INFORMATION SEEKING TOOL BASED ON LEARNING STYLE

NOR LIYANA BINTI MOHD SHUIB

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2013

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **NOR LIYANA BINTI MOHD SHUIB** (I.C/Passport No: **820117075596**)

Registration/Matric No: **WHA090017**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of ~~Project Paper/Research Report/Dissertation~~/Thesis (“this Work”):

INFORMATION SEEKING TOOL BASED ON LEARNING STYLE

Field of Study: **INFORMATION SYSTEM**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date

Subscribed and solemnly declared before,

Witness’s Signature Date

Name:
Designation

ABSTRACT

Learning is enhanced when its activities are aligned with learning style. An integral component of students' learning activities is searching for and retrieving reading materials. However, it is not always easy to find suitable reading materials that match the students' needs. This is mainly because most search engines are built with only content in mind. This study investigates students' learning needs in information seeking, in relation to their learning styles. This involves identifying students' information seeking behaviour from a literature review and preliminary studies. It was found that the reading materials retrieved using current information seeking tools do not match students' learning styles. Thus, this research aims to develop an information seeking tool that considers students' learning style within its retrieval process. The development of such a tool requires the reading materials to be classified based on learning style preferences. Primitive elements, such as text, graphs, and diagrams, have been chosen as identifiers for the above classification. An information seeking tool, with learning style consideration, was then proposed. A prototype was developed to map and match reading materials to students' learning styles. For the evaluation, experiments and surveys were conducted, and the results obtained were then analysed. Our findings show that this new information seeking tool is able to find reading materials that are more closely related to the students' individual learning styles.

ABSTRAK

Pembelajaran dapat dipertingkatkan sekiranya aktiviti pembelajaran yang dilaksanakan sejajar dengan gaya pembelajaran. Komponen penting dalam aktiviti pembelajaran pelajar adalah mencari dan mencapai bahan bacaan. Walau bagaimanapun, mencari bahan bacaan yang sesuai dengan keperluan pelajar bukanlah suatu perkara yang mudah. Ini adalah disebabkan oleh kebanyakan enjin capaian dibina hanya berdasarkan kandungan bahan bacaan. Kajian ini mengkaji keperluan pembelajaran pelajar dalam pencarian maklumat dan kaitannya dengan gaya pembelajaran mereka. Ini melibatkan pengenalpastian pencarian maklumat pelajar daripada kajian literatur dan juga kajian awal. Keputusan menunjukkan yang bahan bacaan yang dicapai menggunakan alat pencarian maklumat yang sedia ada tidak sepadan dengan gaya pembelajaran pelajar. Oleh itu, kajian ini bertujuan untuk membangunkan sebuah alat pencarian maklumat yang mengambil kira gaya pembelajaran pelajar dalam proses capaian. Untuk membangunkan alat tersebut, bahan bacaan perlu diklasifikasikan dengan gaya pembelajaran. Elemen primitif di dalam bahan bacaan seperti teks, graf, dan gambar rajah telah dikenalpasti sebagai elemen untuk diklasifikasi. Sebuah alat pencarian maklumat dengan mengambil kira gaya pembelajaran dicadangkan. Prototaip dibangunkan untuk memeta dan memadankan bahan bacaan ke gaya pembelajaran pelajar. Untuk penilaian, eksperimen dan soal selidik telah dijalankan dan keputusan yang diperoleh dianalisis. Penemuan menunjukkan bahawa alat pencarian maklumat yang dibangunkan mampu untuk mencari bahan-bahan bacaan yang lebih berkait rapat dengan gaya pembelajaran pelajar.

ACKNOWLEDGEMENTS

Alhamdulillah, all the praise goes to Allah for giving me the strength, patience and ability to complete this thesis.

First and foremost, my utmost gratitude to Prof. Dr. Mohd Sapiyan Baba, for his advice, support and guidance throughout this study;

Dr. Rukaini Abdullah, for her ideas and unfailing support as my supervisor;

Mr. Anton Heryanto, for his support and guidance in trying new technology;

My colleagues, for the words of courage and support in various possible ways;

Faculty of Computer Science and Information Technology, for the help and support offered to its postgraduate students;

University of Malaya, for providing resources to carry out this research;

Respondents of this study, for their kind cooperation;

My parent and parent in law, for their prayer, support and encouragement;

And more than anyone else, my husband, Mukhlis and my two children, Ammar and Amani, for their unconditional love and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations	Description
A	Aural Preference
AC	Abstract Conceptualization
AE	Active Experimentation
CE	Concrete Experience
DFD	Data Flow Diagram
ERD	Entity Relationship Diagram
FCSIT	Faculty of Science Computer and Information Technology
FSLSM	Felder and Silverman Learning Style Model
IE	Information Extraction
IEEE	The Institute of Electrical and Electronics Engineers
IST	Information Seeking Tool
K	Kinesthetic Preference
k	Kinesthetic Feature Vector
k-NN	k-Nearest Neighbor
LOM	Learning Object Metadata
LS	Learning Style
LSI	Learning Style Inventory
LSIST	Learning Style Based Information Seeking Tool
M	Mean
MBTI	Myers-Briggs Type Indicator
OPAC	Online Public Access Catalogues
PDF	Portable Document Format

PhD	Doctor of Philosophy
R	Read/write Preference
r	Read/write Feature Vector
RM	Reading Material
RO	Reflective Observation
RQ	Research Question
SPSS	Statistical Package for Social Science
SSADM	Structured Systems Analysis and Design Method
V	Visual Preference
v	Visual Feature Vector
VARK	Visual, Aural, Read/write and Kinesthetic

CHAPTER 1

INTRODUCTION

Learning involves various activities, such as solving problems, group discussions, writing reports, and listening to lectures. One important component of such activities is seeking information. Extra information is needed to supplement what is provided in classes and students need information to resolve problems, make decisions, reduce uncertainties, resolve conflicts, answer questions, and satisfy their curiosity. This helps them understand their courses, enhance their knowledge about the subject, and accelerate their learning.

Students depend on lectures to source good Reading Materials (RMs). RM is one type of learning material. It is the most commonly used learning material and can be easily accessed by students (Commonwealth of Learning, 2005). Several studies have revealed that learning can be enhanced through the presentation of materials that are consistent with a student's particular learning style (Budhu, 2002; Pen˜a et al., 2002; Stash et al., 2004).

Lecturers normally provide RMs for topics covered in a course. Although such RMs are considered suitable for most students taking the course, from our experience, not all students understand what they read. They often find it hard to comprehend the RMs and consequently, they must be able to source alternative RMs to help them learn better. These additional RMs can come from Information Seeking Tools (ISTs) such as internet search engines, digital libraries, or online databases.

1.1 Research Motivation

Learning Styles (LSs) have been shown to enhance learning; particularly when the material matched with students' individual LS (Franzoni & Assar, 2009; Rasmussen, 1998). This is because different students have different preferences in learning, as indicated by their LS. However, some students are disadvantaged because they are either not aware of their LS or do not know their own LS preference (Rogers, 2009). This was equally supported by the results of our preliminary study, which revealed that students have difficulty in finding suitable RM and are unaware of their LS. Without knowing their LS, it is not easy to find the right type of RM. The information may not suit the students' LSs and ability, because of the way it is presented (Yang & Wu, 2009). Students have problems finding suitable RM due to a mismatch of the different attributes of RMs and LSs. Korobili et al. (2011) noted that students have difficulty in finding suitable RMs in their studies. This is in line with (Galvin, 2005) who also noted that students often fail to use appropriate RMs in their research assignments. Therefore, there is a need to align students' attributes such as LS, and RMs' attributes, such as such as the way information is presented; as misalignment will have the most effect on students' understanding of the subject and their attitude in class (Felder, 1995). Students will lose their focus, become demotivated, and think the subject is too difficult; consequently, giving up on the subject (Felder & Silverman, 1988; Renzulli, 1986).

1.2 Problem Statement

With the amount of information and technology that is available today, it is easy to find RM on virtually any topic. Many ISTs can be used to help students find additional RMs. However, they often end up with mountains of information; thus resulting in information overload. As such, existing general ISTs are inadequate in finding RMs based on students' individual preference. A way to solve that problem is to take into account students' LS when developing ISTs and RMs. Personalized ISTs are compared based on the search facilities provided. To date, no tools are available that can retrieve RMs based on students' LS (Apriyani & Hasibuan, 2008; Franzoni & Assar, 2009; Savic & Konjovic, 2009; Yang & Wu, 2009; Hassan et al., 2010; Bachari et al., 2011; Klašnja-Milićević et al., 2011). Currently, RMs are not categorized according to LS and there are no guidelines that can be used to categorize RM based on LS. Hence a method to classify and map RM to LS is required and an IST that integrates LS in its search query needs to be developed to retrieve RMs based on students' LS.

1.3 Aims and Objectives

The primary goal of this research is to investigate how students' learning needs for RM can be accommodated based on their preferred LSs. More specifically, this research aims at developing an IST that can retrieve RMs that matched students' LS. To accomplish the above, the research has the following objectives:

1. To investigate the need for an IST with LS consideration
2. To determine a method to categorize RMs based on LS
3. To determine a method to map RM onto students' LS preference
4. To develop and evaluate the LS based IST prototype

1.4 Research Questions

The research questions (RQ) for each objective are presented below,

Objective 1: To investigate the need for an IST with LS consideration.

RQ 1: Are students having difficulties to find suitable RM for their learning needs?

RQ 2: Is there any relation between RM and LS?

RQ 3: Are students aware of their LS?

RQ 4: Are there available ISTs that support students' LS?

Objective 2: To determine a method to categorize RMs based on LS.

RQ 5: How can primitive elements in RM be categorized based on LS?

Objective 3: To determine a method to map RM onto students' LS preference.

RQ 6: How can RM be mapped onto students' LS preference?

Objective 4: To develop and evaluate the LS based IST prototype.

RQ 7: Can the prototype perform better?

1.5 Research Design

This research consists of three steps which are Identification of the Problem, Design and Development of LSIST and Evaluation of LSIST as shown in Figure 1.1. In the first step, the problem is identified from literature review and preliminary studies. Then, the architecture for the *Learning Style based Information Seeking Tool* (LSIST) was proposed. Primitive elements in RMs, such as words, figures, charts, and diagrams were categorized based on LS. Algorithms for extracting these primitive elements and mapping RMs to students' LS were developed. An experiment to evaluate the prototype was then conducted and the findings were discussed. With this tool, RMs could be retrieved based on students' preferred LSs.

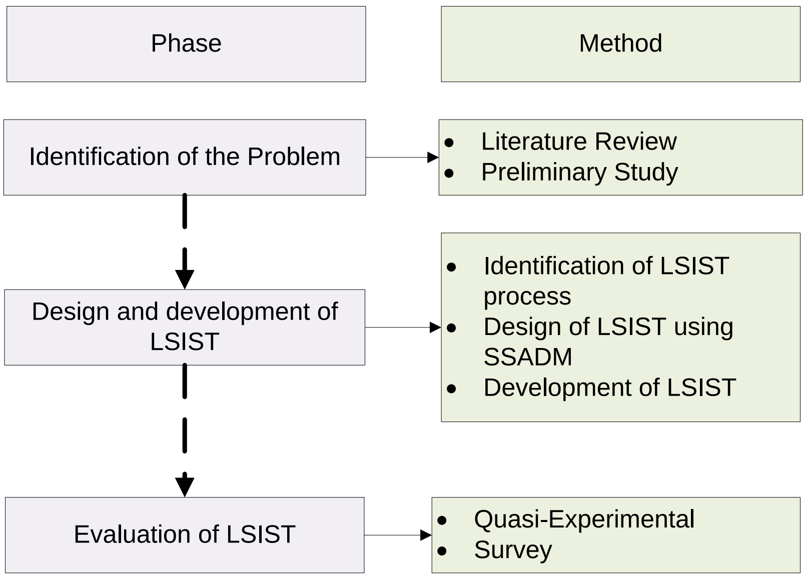


Figure 1.1: Research Design

Step 1: Identification of the Problem

The research problem is identified by reviewing previous works related to students' difficulties in finding suitable RM, information seeking, ISTs, information extraction and the concept of LSs. To confirm the findings from literature, preliminary studies are then conducted to investigate students' information seeking behavior and LS awareness.

The preliminary studies consist of:

- **Questionnaire 1:** A Survey on Students' Information Seeking Behaviour

  The objectives are to study the information seeking behaviour of Computer Science students and to ascertain the problems faced by them in seeking information.

- **Questionnaire 2:** A Survey on Students' LS Awareness

  The objectives of the survey are to investigate students' awareness of LS and to identify students' LS using VARK LS test.

Step 2: Development of LSIST

This process consists of four tasks:

1. **Classification of Primitive Elements in RM**

   The identification and classification of primitive elements in RM are discussed in section 5.3.1 and 5.3.2.

2. **Mapping of RM onto LS Preference**

   The classified primitive elements in RM are then mapped onto LS preference (See section 5.3.3). Each primitive element is standardized using identifiers from Learning Object Metadata (LOM). LOM is discussed in section 2.2. The extraction of the identifiers in RM and its calculation using *Feature Extraction* algorithm is discussed in section 5.4.3.

3. **Matching of RM onto Students' LS Preference**

   An algorithm to map RM onto student' LS preference is presented in section 5.4.5. It uses k-Nearest Neighbor classification model to match students to the appropriate RM based on their LS preference.

4. **Design and Implementation of LSIST**

   The architecture of LSIST is proposed in section 6.1. A prototype based on the architecture is designed and developed using Structured Systems Analysis and Design Method (SSADM) (Weaver et al., 1998). Design and development of LSIST is discussed in section 6.2.

Step 3: Evaluation of LSIST

The evaluation of the prototype is done via an experiment and survey as below.

1. **Experiment**

   The evaluation for the prototype is conducted using Quasi-Experiment design: Pretest/Posttest, Non-equivalent Multiple Group Design. Respondents evaluate RMs retrieved from *Keyword based Search* (Pretest) and *LS based Search* (Posttest). The results are discussed in section 7.2.1.

2. **Survey**

   A survey is conducted to get feedback from respondents about the prototype. The results are discussed in section 7.2.2.

1.6 Research Scope

This study was conducted based on certain delimitation to focus the research. The boundaries are:

- Only Computer Science postgraduate students are selected for investigation. Hence, the collection of research material is confined to those in Computer Science
- Research materials only involve RM. Videos and audios are not included
- The RM used is of the Portable Document Format (PDF) type
- Does not include image or symbol processing

1.7 Contribution of the Research

The contributions of this research to the current literature are as follows:

- Learning Style Based Information Seeking Tool (LSIST) Architecture
- Identification and classification of primitive elements in RM according to LS. Identification of primitive elements in RM is done by using iText and then classified according to *Text* and *Non-Text* categories. The classified primitive elements are then mapped onto LS preferences using identifiers from Learning Resource Type in LOM.
- *Feature Extraction Algorithm*, a novel algorithm that extracts and classifies RM onto LS preference. Identifiers of primitive elements in RM are extracted and calculated using *Feature Extraction Algorithm*.

- *Matching Algorithm*, a novel algorithm that match RM onto students' LS preference. RM was mapped onto students' LS preference using k-NN classification method.
- LSIST that retrieves RM based on students' LS. The prototype of LSIST is developed and evaluated using quasi-experimental and a survey.

1.8 Significance of the Research

Currently, no IST exists that considers student's LS in its retrieval process. The incorporation of LS in IST is a new method for IST development. Processes such as the classification of primitive elements, the mapping of RM to LS, and the mapping of RM to students' LS, have not been done elsewhere. This research introduces a new focus in this area which can be expanded further.

1.9 Thesis Overview

Finding suitable RMs that match students' individual preference during information seeking is important. The main aim of this research is to develop an IST based on students' LS. Thus, this research investigates and presents students' learning needs and their difficulties in information seeking process. Chapter 2 examines the capabilities of existing ISTs and discusses the concepts of LS. Chapter 3 explains the stages involved in this research.

Chapter 4 reports on the analysis of the results from two surveys conducted. Particular interest is paid to answering questions such as:

- Are students having difficulties to find suitable RM for their learning needs?
- Are students aware of their LS?

Chapter 5 discusses how RM can be mapped and matched onto LS preference. The prototype design and development is described in Chapter 6 followed by its evaluation in Chapter 7. Chapter 8 concludes the thesis.

CHAPTER 2

INFORMATION SEEKING FOR LEARNING

2.1 Students' Learning Needs

Every student has different learning needs. Learning needs is something that can be identified from a learning journey to improve or enhance existing skills or knowledge. Identifying learning needs is important to make learning better. If students understand their learning needs and they can find suitable material that meet their need, it will improve their learning.

To satisfy students' learning needs, the following three factors must be taken into consideration (Tomlinson, 2009):

- Readiness

  This refers to students' knowledge and skill level
- Interest

  This refers to topics that students want to explore or motivate the student
- Learning profile

  It consists of learning style, culture, gender, intelligence preference and environment preference

Students need to select learning materials that fulfill the above factors to maximize their learning outcomes. There are various types of learning material that are prepared based on different level of lesson content and different type of instructional material. Lesson content depends on students' readiness level while instructional materials depend on students' learning preference. During learning process, students are aware of their readiness level based on their learning syllabus, year of study and examination grade that they are getting. So they usually know their readiness level.

However, choosing instructional material that suits their learning needs is not an easy task**.** Instructional material can be defined as the physical means via which instruction is presented to the students (Reiser & Dempsey, 2007). This includes:

- Reading materials: Printed and electronic text (book, textbook, manual, journal)
- Audio materials: Videos, audiotapes
- Visual materials: Real objects and model, maps, still pictures, slides
- Computer-mediated materials: Interactive whiteboard, interactive video, multimedia

Instructional materials are prepared and provided by educators. Students usually have a limited access of the materials and most of them can only be used with educators' permission or supervision. The only material that can be accessed easily is reading materials.

2.2 Reading Material

Reading materials (RMs) are written document to be read. They contain data that are organized in the form of meaningful information. Students need information to carry out activities such as resolving problems, making decisions, reducing uncertainties, resolving conflicts, answering questions and satisfying curiosities. It helps them understand their courses.

RMs can be obtained from different sources: primary, secondary and tertiary information sources (Omar & Mansor, 2005):

- Primary information is collected direct from the original source without any editing, interpretation or evaluation. Examples are diaries, autobiographies, interviews, survey, letter, blogs, and minute of meeting. This type of source is the closest to the actual event, time period and individual in question. However, primary information is hard to acquire.
- Secondary information is interpreted from primary information, hence more available and easier to use. However, it is less reliable since there is a possibility of misinterpretation from the secondary sources. This type of information is usually critically evaluated to ensure its integrity. Examples of such sources are books, journals, technical reports, conference papers, thesis and dissertations, and handbooks.
- Tertiary information act as tools for understanding and locating information. This type of information sources functions as a guide to the real information and makes information seeking easy and practical. Examples are bibliography, index and dictionaries.

With the help of technology, most of the RM especially in education domain can now be accessed and easily found online. Many organizations have digitized the printed RMs. Publishers also published RM in print and online forms, making the information more accessible.

RM is a learning object. Learning object is defined as digitized entity which can be used, reused or referenced during technology supported learning (Rehak & Mason, 2003). The Institute of Electrical and Electronics Engineers (IEEE) introduced IEEE 1484.12.1 - 2002 Standard Learning Object Metadata to facilitate search, evaluation, acquisition, and use of learning objects (IEEE, 2002).

Learning Object Metadata (LOM) is used to describe a learning object to support learning. Relevant attributes of learning objects such as type of object, author, owner, terms of distribution, and format are described. One of the attributes that groups the educational and pedagogic characteristics of the learning object is Learning Resource Type. Learning Resource Type identifies specific kind of learning resource such as exercise, simulation, questionnaire, diagram, figure, graph, index, slide, table, narrative text, exam, experiment, problem statement, self assessment, and lecture (IEEE, 2002).

2.2.1 Categorization of Reading Material

RM contains primitive elements such as characters, figures, tables and diagrams (Aumann et al., 2006). RM can be presented in various combinations of primitive elements. Scholarly RMs like scientific and technical RMs do not only contain text but also tables, pictures, flow charts, drawings, maps, figures and mathematical expressions (Antoine et al., 1992).

RM can be categorized by classifying their primitive elements into categories. Dori et al. (1997) classify primitive elements in RM as text and non-text while Aumann et al. (2006) categorize primitive elements in RM into three categories which are textual, image and graphic line. Text category consists of letters alphabet, character and symbols and can be presented as words in lines, columns and paragraphs (Bayer et al., 1992) while non-text category is commonly referred to as graphic and labeled as a figure within the embedding RM.

Logical text can be used to extract *Non-text* primitive elements. Logical text is text information that is associated with one or more images that fit on one page (Christodoulakis & Theodoridou, 1986). It provides hints on the semantics of the image. For example captions of the images, the text surrounding the images in RM and the text reference to the images. Logical text is widely used to represent image in information extraction research (Lu et al., 2006).

These various forms of primitive elements illustrate and explain contents in the RM (Dori et al., 1997). In many cases, graphics provide the basis for the information conveyed by the RM and used to illustrate the key ideas (Tombre & Dori., 1997). Therefore, appropriate combination of primitive elements is important because it can enhance understanding and help students encode the information more effectively (Lu et al., 2006).

2.2.2 Extracting Information from Reading Material

In order to extract information from RM, information extraction technology is used. Information extraction (IE) is a technique used to analyze and locate specific information from RMs (Kaiser & Miksch, 2005). Most of IE techniques are based on the extraction of the textual content (Aumann et al., 2006).

There are a few techniques that have been used for IE (Fong et al., 2002):

- Lexicon-based

  Information is extracted from contents identifiable by a lexicon or list of root words, including individual keywords and phrasal units, proper names, abbreviations, numbers, and codes that commonly appear in a particular information domain such as FRUMP (DeJong, 1982).
- Syntax-based

  Relies on data presented in certain predefined formats, or when the document structure is marked up using a set of predefined tags such as Jedi (Huck et al., 1998).
- Heuristic-based

  Additional rules are used to improve the extraction, e.g. common conceptual knowledge and regular patterns observed from data sources results.
- Machine learning-based

  Algorithms developed in natural language processing.

However none of these techniques can be used with Portable Document Format (PDF) (Hassan & Baumgartner, 2005). A PDF document is composed of text, graphics, table and images (Déjean & Meunier, 2006; Yuan et al., 2006). It is one of the most popular formats and widely accepted nowadays because of the wrapping ability of the PDF document and it restricts the direct editing of contents which makes them more secure (Baker et al., 2008; Pitale & Sharma, 2011).

There are several works on extraction of object in PDF documents. One of them is extracting text from PDF (Yuan et al., 2006). This technique transforms the textual information parsed from PDF files into semi-structured information by injecting additional uniform tags guided by an extensible rule set. While Yuan et al. (2006) extract text, Baker et al. (2008) proposed a technique to extract mathematical expressions from PDF documents. They achieve this by extracting information using fonts and geometric positions of single characters from the PDF source. Lu et al. (2009), Hassan (2009) and Shao and Futrelle (2006) focuses on extracting graphics and image from PDF documents. However, none of them consider LS in the extraction process.

Recall and precision is the most frequent measure used when referring to information extraction and information retrieval (Kaiser & Miksch, 2005). Raghavan et al. (1989) defined recall as the ratio of the number of relevant documents that are retrieved to the total number of relevant documents and precision as the number of relevant documents retrieved divided by the number of retrieved documents.

To extract PDF content, PDF are extracted into a format that can be manipulated using PDF extraction tool (Shao & Futrelle, 2006). There are five extraction tools that have the capabilities to extract image, text and font which are PDFL1b TET, 3-Height PDF Extract, jPadel, ICEpdf and iText (Pitale & Sharma, 2011). iText is the best tool because it have all the capabilities as others and free of cost. Others are commercial tools.

2.3 Information Seeking

The role of information seeking in education has increased dramatically over the last few decades. This is mainly due to the powerful development of new information and communication technologies and an increasingly student-centred and problem based pedagogical orientation. In order to handle this, students need the ability to critically seek, evaluate, and use information, and tools for information retrieval. Limberg (1999) identified a close interaction between students' understanding of the topic of their task and their approach to seeking information.

With changes to information accessibility, students need to know how to seek RMs. As learners, students are not satisfied with just locating relevant or useful information. Their goal (in information seeking) is no longer to merely get relevant information, but to get information that leads to a new understanding in the process of learning (Kuhlthau, 1997).

Students need information to seek answers, reduce uncertainty, and make sense of something in their learning process. They need suitable and sufficient information to help them understand their course effectively. A student’s good knowledge of information seeking can help them find suitable information to fulfil their information needs better.

Wilson (1999) defined information seeking as a purposive seeking of information and a consequence of a need to satisfy some goals. Meanwhile, Marchionini (1995) defined it as a process in which humans purposely engage in changing their state of knowledge. Wilson’s definition is more towards fulfilling information needs, whilst Marchionini’s is more towards learning and problem solving. Hence, information seeking can be defined as a process of finding information with a specific information need, in order to improve the level of understanding or knowledge related to a problem.

In seeking information, a student employs a particular information searching behaviour, which is the micro-level behaviour that a searcher uses in interacting with all types of information systems (Wilson, 2000). It includes all interactions with systems, such as human computer interaction, intellectual and mental acts i.e., adopting a Boolean search strategy, and evaluating the relevance of information retrieved. In this study, the use of the term *information seeking* implies the employment of *information searching*.

Rapid developments of information environments have brought significant changes to how information seekers look for information. Nowadays, users; especially the Millennial Generation (i.e., those born between 1979 and 2000 - also dubbed as the Net Generation), adopt convenience as a criterion in choosing information sources or strategies. Users who adopt convenience information seeking use minimal amounts of effort to satisfy their needs; this has been delineated as the Principle of Least Effort, as first proposed by Zipf (1949). This principle has been used as an underlying notion in convenience information seeking, which demonstrates user’s preference to use easier information resources, while seeking for information (Connaway et al., 2011; Liu & Yang, 2004). However, “satisficing” is another intriguing phenomenon that reflects the Principle of Least Effort, which was defined by Simon (1955). Satisficing, which is a combination of the words ‘satisfy’ and ‘suffice’, reflects upon the user’s behaviour to choose what is satisfactory (or good enough) rather than what is best (Byron, 2004). Satisficing is characterized by the amount of effort that the information seeker is willing to make in finding information. In this case, information quality is normally compromised in favour of the most convenient method.

Students access and use various information retrieval tool or Information Seeking Tools (IST) for diverse reasons (Cooke, 2001). Various types of IST are available to students.

For the purpose of this study, four types of IST were used, namely: internet search engines (e.g. Google, Yahoo, and AltaVista); Online Public Access Catalogues (OPAC); Online Databases (e.g. Elsevier, IEEE, and ACM), and Digital Libraries, such as Dspace@UM and DigiLibraries.com. This will be discussed further in Section 2.3.1.

A review of information seeking studies by Catalano (2013) reported that none of the studies focused on computer science graduates. Although Saad and Zainab (2007; 2009) conducted studies on how undergraduates of computer science searched for information; they could not be used to measure graduates of computer science, because their prior knowledge was different.

Little is known about how computer science graduate students behave when seeking information. Researchers were generally more interested in studying novice user's behaviour than studying users with an Information and Communication Technology (ICT) background similar to that shown by computer science graduate students, as stated by Ismail et al. (2009).

Since computer science graduate students have received little empirical attention from researchers, their information seeking behaviour often resembled that of other disciplines or undergraduates. With higher levels of education and an ICT background, computer science graduate students possibly have a different information seeking behaviour. This assumption is based on the premise that computer science graduate students have relatively more Information Technology (IT) and information seeking skills, due to their education background, experience, and skill in ICT. As graduate students, their engagement in the research process would be more proficient than undergraduate students, as articulated by Catalano (2010).

This apparent gap suggests that it is necessary to explore their information seeking behaviours and investigate to what extent they might constitute being a unique user group. Hence, computer science graduate students were chosen to present a unique user group, due to their relatively high levels of IT, research, and information seeking skills, as compared to others.

2.3.1 Information Seeking Model

Various information seeking models were developed. The following models are generally the most widely used in research. Many researchers develop their own phases based on one of these models. In 1989, Ellis proposed a model employed from Gaser and Strauss's grounded theory. In his model, he uses the term *features* rather than *stages*. These features are named and defined as follows:

1. Starting - begin seeking information
2. Chaining - following footnotes and citations in known material or 'forward' chaining from known items through citation indexes
3. Browsing - 'semi-directed or semi-structured searching'
4. Differentiating - using known differences in information sources as a way of filtering the amount of information obtained
5. Monitoring - keeping up-to-date or current awareness searching
6. Extracting - selectively identifying relevant material in an information source
7. Verifying - checking the accuracy of information
8. Ending - may be defined as *tying up loose ends* through a final search.

In 1991, Kuhlthau published her Information Search Process (ISP) model, based on a major research of information users, based on psychology, education, and information science. Kuhlthau (1997) identified six steps that students were engaged in when conducting a research task, namely **i**nitiation**, s**election**, e**xploration**, c**ollection**, s**earch closure, and presentation.

The strength of both models is that they are based on empirical research and have been tested in subsequent studies (Wilson, 1999). Ellis (1989) and Kuhlthau's (1991) models can be applied to any domain; however, their models make no claim to consider many of the factors and variables generally considered in information seeking research i.e., the type of need and sort of information or other "help" might satisfy it; or the availability of sources and their characteristics (Case, 2007).

In this research, we focus on information seeking behaviour in an online environment. Therefore, the Marchionini model (1995) is deemed most suitable, because it was designed as a generic construct to describe information seeking processes from any electronic information source. Marchionini's (1995) model focuses specifically on the information search processes of electronic system users, and their actual information seeking activities. Marchionini's information seeking model is shown in Figure 2.1

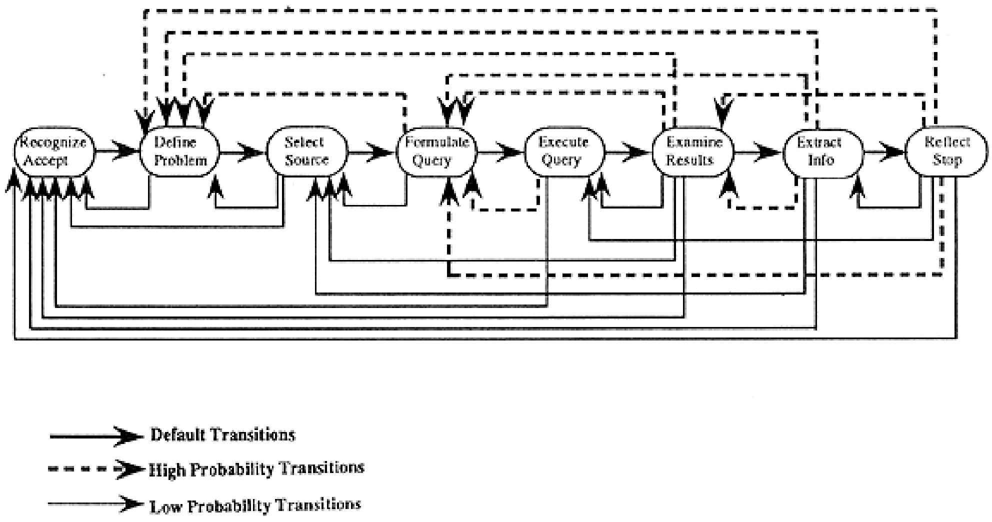


Figure 2.1: Marchioninis' Information Seeking Model (Marchionini, 1995)

2.3.2 Information Seeking Process

Information seeking involves a series of process; analysis of information need, selection of information retrieval tool or IST, selection of search strategy, evaluation of the results and iterating the steps if necessary. This is based on Marchioninis' Information Seeking model (Marchionini, 1995; 1998). Each step must be carried out properly to maximize the information seeking results. The whole process is shown in Figure 2.2.

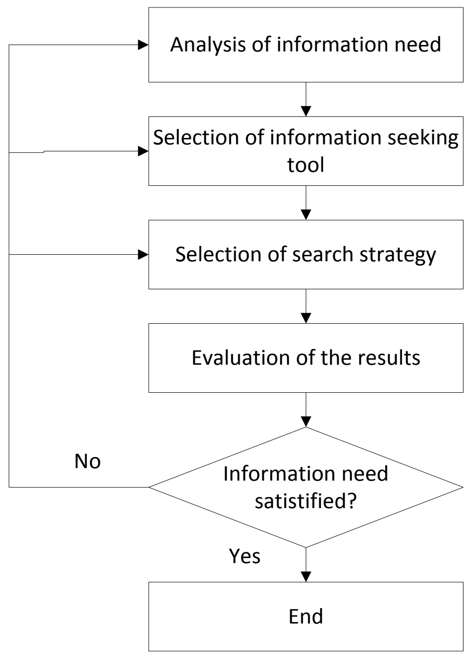


Figure 2.2: Information Seeking Process

Step 1: Analysis of information need

Information need is defined as the recognition that one's knowledge is inadequate to satisfy a goal that one has (Atkin, 1973). The goal can be in the form of seeking an answer to a question or making sense of something in the learning process. It is important to establish as precisely as possible the information needs before starting to find the information. The need must be identified in terms of keywords. Failure to formulate the correct keywords leads to search failure. For example; suppose a student wants to find information about *information seeking* but he only enter keywords such as *information* or *seeking information;* the results may not be suitable for his information needs.

Step 2: Selection of IST

Once information need has been analyzed and understood, students can then select an IST to begin the searching. IST is a system that informs the user about the existence or non-existence of the document related to the user request (Lancaster, 1968). It is developed to facilitate information seeking. IST helps student to find information by organizing the information in order to make the information easier to search and identify. There are many IST available on the Internet. Students can choose available tools to retrieve the desired materials. Students may access and use these various ISTs for diverse reason (Cooke, 2001). Examples of IST are Online Public Access Catalogues (OPAC), Internet search engine, subject directory, online database, and digital library.

Step 3: Selection of search strategies

A search strategy is formulated to retrieve information containing the selected search terms identified from the analysis of information needs. Search strategy depends on the search function employed by the IST itself. All existing search strategies are based on comparison between the query and the stored documents. Different search strategies produce different results. Examples of basic search strategies are Boolean search, truncation, proximity search, field or meta tag search, limiting search, range search and advanced search techniques such as cluster representatives, relevance feedback, search indexes, and controlled vocabulary. Common search strategies are summarized in Table 2.1.

Step 4: Evaluations of the results

An evaluation of the results is crucial because when wrong information is chosen then the information need is not fulfilled. In evaluating the search results, the information found need to be verified to ensure its relevance to the information need. This is important to guarantee that the information found is useful and understandable.

Table 2.1: Search Strategies

Search Strategies	Description
Boolean Search	Boolean search is the most commonly used search strategies that combine search term with Boolean operators, such as *AND, OR*, and *NOT.* The operators are used to refine and extend search query. The terms and operators can be combined using parentheses. For example, the search term 'Digital AND Library' retrieves all records that contain both the terms.
Truncation	Truncation is a search facility that enables a search to be conducted for all the different forms of a word having the same common root. There are a few types of truncation such as left-hand truncation, right-hand truncation and character masking or wild cards. For example, the search term Librar* retrieves all records that contain Library, Libraries, Librarian and so on.
Proximity Search	Proximity search allows users to define the distance between keywords. It retrieves documents that have two or more separately matching term within a defined distance. Distance is the number of intermediate words. It is quite similar to Boolean search but more specific and restricted. It can be used with other query operators like *NEAR*, *NOT NEAR, FOLLOWED BY* and *NOT FOLLOWED BY*. For example, the search term Library *NEAR* Digital retrieves all records that contain Digital Library, Library of Digital Object and so on. By using this type of search, it avoids documents where the words are spread across a page or in unrelated articles.
Field or meta tag search	Field or meta tag search is used to restrict a search to a specific field to obtain more precise results. A user can type a field search such as *CONTAIN* before the keyword or selecting an option such as *in the entire document*, in the title of the document, *in the author's name*, and *in the assigned keywords*. For example the search term Library with *in the title of the document* option retrieves all records that contain Library in the title of the records.
Limiting Search	Limiting search is used to restrict a result using certain criteria such as date of publication, language, type of collection, type of material and year of publication. The criteria depend on the IR system. This type of search is very useful to limit results especially in the web environment. For example the search term *Library with year of publication option in 1992* retrieves all records that contain Library in 1992 year of publications.
Range Search	Range search is used with numerical information to select records within certain ranges. It is used with operators such as Greater than (>), Less than (<), Equal to (=), and Not equal to, etc (/=) or (<>).

This information seeking process is finished if the information need is fulfilled. Otherwise, the process is iterated by changing keywords, ISTs or search strategies used until the desired information is found.

2.4 Difficulties in Seeking Reading Material

With the amount of information and the technology that is available today, it is easier to find information on any topic. However, the information obtained may not necessarily help students in their learning process. Information in this case can be defined as RM that is used for learning. Galvin (2005) noted that students often fail to use appropriate RMs in their research assignment.

One of the problems is information overload. For example, when a student seeks for information on a particular subject, he looks at all possible information sources such as books, articles, and journals and ends up with a mountain of information, making it difficult to choose the most relevant and appropriate sources (Cooke, 2001; Iverson, 2005).

Students do not necessarily find information that suits their style of learning. They usually evaluate information based on the keyword that they used. As long as the content of the information have the keywords, they tend to assume that it is the right information for them. Such students need help from their lecturers to find suitable RMs.

Lecturers can help the students find the information on the topic of their study at the appropriate level of knowledge but not all students can understand the same form of information. This situation is called *one size fits all*. For a standard learning process, lecturers normally provide the same form of information for all students. Students have to adapt themselves to the information given. However, this does not work in practice. They end up not comprehending the information given.

Furthermore, RMs available are not categorized according to the way information is presented. This does not help students choose the right RM. It is difficult or even impossible for student to learn when the material is not presented in their preferred style (Gregorc, 1985). Even though the information retrieved is relevant to their topic and level of knowledge; it does not contribute to better understanding of the subject.

2.5 Need for Learning Styles

In 1983, Gardner proposed the Multiple Intelligence Theory. It has been widely used in education, teaching, and training communities. Gardner identified eight different kinds of intelligence that reflect different ways of interacting, namely musical - rhythmic, visual - spatial, verbal - linguistic, logical - mathematical, bodily - kinesthetic, interpersonal, intrapersonal, and naturalistic. Figure 2.3 shows these intelligences.

Figure 2.3: Multiple Intelligence (Gardner, 1993)

According to Gardner, each individual has a unique combination, or profile (Gardner, 1983). Even though this theory suggests different kinds of intelligence, Gardner suggests that each individual has a different configuration of intelligence i.e., none of us are the same. This is influenced by individual characteristics, such as prior knowledge, education level, past experience, level of literacy, motivation, task confidence, aptitude, and learning style (Sadler-Smith, 1996; Howles, 2006). These differences affect learning activities and outcomes. One of the most important factors that affect the outcomes of learning is learning style (Felder, 1995).

Learning Style (LS) is important in a learning process. It can help students enhance their learning capabilities. For example, if students know and understand their LS, they can choose RMs that are suited to their LS, which can assist them in understanding the subject better. This could enhance their learning capabilities.

LS represents several inter-related elements, namely modalities, instructional preferences, and learning strategies. The use of these elements in learning can be influenced by the type of learning task and the environment (Peterson, 2009). For the purpose of this study, modality is used. Therefore, LS is defined as the preference (or predisposition) of an individual to perceive and process information in a particular way, or combination of ways (Zapalska & Brozik, 2006).

2.5.1 Learning Style Models

Research on LSs emerged in the early 1960's (Howles, 2006). There are large volumes of LS model in the literature. Curry (1983) had classified them into four categories according to different aspects of learning preference. Learning preferences may be defined as the favoring of one particular mode of learning over another. The four categories are:

1. Personality

   Influence of basic personality traits on acquiring and integrating information
2. Information Processing

   Intellectual approach to assimilate information
3. Social Interactions

   Social interaction between individual
4. Instructional Preference

   Individual preferred environment for learning

LS models in each category will be discussed briefly below.

1. Personality Category

This category represents LS models that have influence of basic personality traits on acquiring and integrating information. These models control central personality dimension which is the cognitive personality style that addresses an individual’s approach to adapt and assimilate information (Curry, 1983). There are three well known models that represent personality models as described below.

Witkin's LS Model

Witkin's LS model (Witkin et al., 1954) construct of field dependence and field independence is based on individual’s ability to extract details from a context. It can be measured using the Group Embedded Figures Test (Witkin et al., 1971). This approach is related more to ability than style (Sternberg & Grigorenko, 1997).

Myers-Briggs Type Indicator

Myers-Briggs Type Indicator (MBTI) is based on Karl Jung's theory of psychological types. It is used to characterize people according to sixteen types which are the combination of eight scales. Its dichotomous scales measure extroversion versus introversion, sensing versus intuition, thinking versus feeling, and judging versus perception. Even though this model is accepted widely, it examines personality types rather than LSs (Atkin et al., 2001).

Felder and Silverman LS Model

The Felder and Silverman LS model (FSLSM) categorizes students using five spectrums which are sensing versus intuitive, visual versus verbal, inductive versus deductive, active versus reflective and sequential versus global (Felder & Silverman, 1988; Felder, 1995). FSLSM is measured using Index of Learning Style (Graf et al., 2007).

2. Information Processing Category

This category represents models that have intellectual approach to assimilate information. These models focus on the processes by which information is obtained, sorted, stored and utilized. They are more adaptable for learning strategies. Examples of such models are Gregorc's Style Delineator (Gregorc, 1982), Kolb's LS Inventory (Kolb, 1984), McCarthy 4Mat Curriculum Development Model (McCarthy, 1987), Honey and Mumford (Honey & Mumford, 1992), Gardner Multiple Intelligence (Gardner, 1993) and VARK model (Fleming & Mills, 1992). Most of the models in this group are measured through self-report and self-assessed instruments.

Gregorc's Style Delineator

Gregorc (1982) proposes that people differ in the way they organize space and time. Gregorc's Style Delineator classifies individuals into two dimensions: firstly, based on the way information is comprehended (abstract or concrete); and secondly, based on the way information is arranged (sequential or random). Then this model categorizes individual into four types; abstract sequential, concrete sequential, abstract random and concrete random.

Kolb's LS Inventory

Kolb's LS Inventory is based on the Experiential Learning Theory (Kolb, 1984;1999). His model outlined two levels. The first level is a four-stage cycle which is Concrete Experience (CE), Reflective Observation (RO), Abstract Conceptualization (AC) and Active Experimentation (AE). The ideal learning process engages all stages in response to situational demands. In order for learning to be effective, all stages must be incorporated. The second level represents the combinations of two preferred stages as illustrated in Figure 2.4. The LSs are as follows:

- **Diverging (CE/RO)**

  Divergers tend toward concrete experience and reflective observation. They are imaginative and are good at coming up with ideas and seeing things from different perspectives

- **Assimilating (AC/RO)**

  Assimilators are characterized by abstract conceptualization and reflective observation. They are capable of creating theoretical models by means of inductive reasoning

- **Converging (AC/AE)**

  Converges are characterized by abstract conceptualization and active experimentation. They are good at making practical applications of ideas and using deductive reasoning to solve problems.

- **Accommodating (CE/AE)**

  Accommodators use concrete experience and active experimentation. They are good at actively engaging with the world and doing things instead of merely reading about and studying them.

Kolb's LS model is measured using LSs Inventory (LSI). The Kolb's LS model has been used as a basis to develop other LS inventories like 4MAT System by McCarthy (McCarthy, 1987) and Honey and Mumford model (Honey & Mumford, 1992).

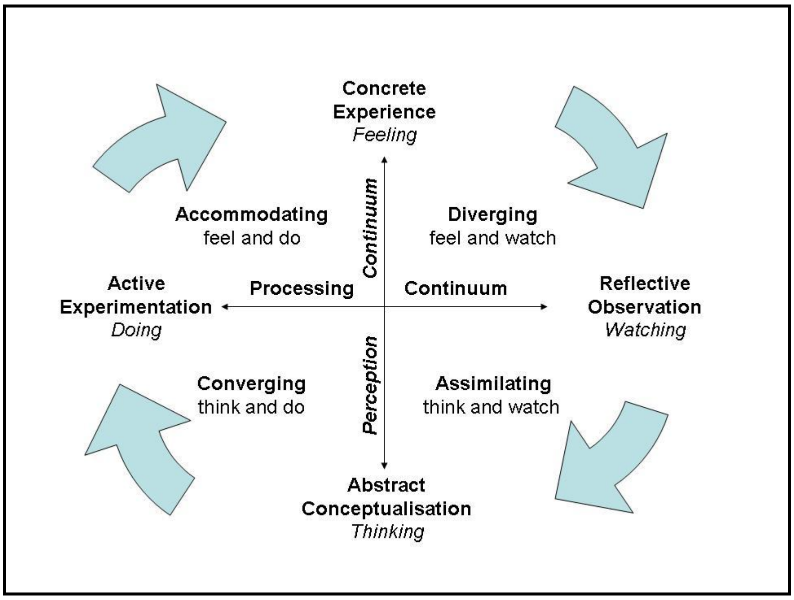


Figure 2.4: Kolb's LS (Kolb, 1999)

McCarthy 4Mat Curriculum Development Model

McCarthy's 4MAT Curriculum Development Model (McCarthy, 1987) attempts to integrate LSs and hemispheric processing. This model assesses preferences for using the left or right hemisphere of the brain. Like Kolb, it is also based around a four stage learning cycle and offers four types of LS which are innovative, analytic, common sense and dynamic.

Honey and Mumford Model

The Honey and Mumford model (Honey & Mumford, 1992) was developed from Kolb's LS model for use in commerce. This model is also based on learning cycle. This leads to the identity of four types of LS which are activists, reflectors, theorists and pragmatists.

Gardner Multiple Intelligence Model

This model is based on the Gardner's' theory of multiple intelligences (1993). In this model, he suggests that each individual has eight distinct areas of intelligence, namely spatial, linguistic, logical-mathematical, bodily kinesthetic, musical, interpersonal, intrapersonal, and naturalistic (Gardner, 1993; 1999).

VARK Model

Visual, Aural, Read/Write and Kinesthetic (VARK) (Fleming & Mills, 1992) was also derived from Gardner's' theory of multiple intelligences (1993). VARK has the following four preferences (Fleming & Baume, 2006):

- **Visual (V)**

Visual learner refers to those who prefer the use of diagrams, pictures, videos, slides, graphs and flowcharts to represent printed information.

- **Aural (A)**

Aural describes a preference for information that is either heard or spoken.

- **Read/write (R)**

This preference categorizes those who prefer printed words and text as a means of taking in information

- **Kinesthetic (K)**

The Kinesthetic preference refers to learning achieved through experience and practice.

No individual is restricted to only one of the four preferences: V, A, R, or K. People usually exhibit a strong preference for one particular preference, whilst having a relative weakness or strength in some of the others. One other preference is Multimodal. Multimodal refers to having a strong preference for more than two preferences.

While most of the models above discuss individual differences based on learning activity preference, the VARK model introduces a LS model that uses sensory modality. Sensory modality is a combination of perception and memory i.e., how the mind receives and stores information. Individuals have different sensory modality preferences when internalizing information. Sensory modality is one of the more practical and recently popular ways to define and assess LS that one prefers when learning (Dobson, 2009). Even though the credibility of this model is questionable, VARK has been tested and validated empirically by Leite et al (2009).

The VARK LS model is chosen because:

- VARK questions are based on real-life situations, so that respondents can easily relate to them (Rogers, 2009)
- Respondents can select more than one answer that suits them for each question; thus allowing them to identify whether they possess multiple modes of learning (Slater et al., 2007).

- Four LS preferences outlined in VARK relate well to the eight intelligence dispositions identified by Gardner and outlined by Silver (Gardner, 1999; Silver et al., 2000).
- It is an accessible method to understand and explain peoples' preferred ways to learn.

3. Social Interaction Category

Social interaction category deal with how students interact with their peers in the classroom. This category includes Grasha and Reichmann (Grasha & Reichmann, 1974) and Dunn and Dunn LS Inventory (Dunn & Dunn, 1978).

Grasha and Reichmann Model

Grasha and Reichmann model sorts' individuals on the basis of six LSs which are independent, dependent, participant, avoidant, collaborative, and competitive (Grasha & Reichmann, 1974).

Dunn and Dunn LS Inventory

This LS model assesses individuals in terms of global approaches to learning (Dunn & Dunn, 1978). This consists of environmental, emotional support, sociological composition, physiological and psychological elements.

4. Instructional Preference Category

This category includes instructional preferences that are affected by environmental conditions (temperature, light, sound, seating comfort); emotional issues (motivation, persistence, and responsibility for structure and direction in learning activities); sociability (preference to work alone, in pairs or small groups, or with an authority); and the varying contexts of learning. An example of this category is Processing Model (Keefe, 1989). This model addresses the individual's preferred environment for learning. Student obviously has little control at this level (Curry, 1983).

2.5.2 Learning Style Model Suitability for Reading Material Context

We have discussed briefly LS models in section 2.5.1. Each type of the LS model serves different aspects of learning. Table 2.2 summarizes the LS models with respect to learning.

Table 2.2: LS Models

LS Model	Personality Traits	Individual Abilities	Learning Activities Preference	Learning Approaches Preference	Sensory Modality
Dunn and Dunn LS Inventory				x	
FSLSM	x			x	
Gardner Multiple Intelligence	x	x			
Grasha & Reichmann				x	
Gregorc's Style Delineator		x	x		
Honey & Mumford			x		
Kolb's LS Inventory			x		
McCarthy 4Mat Curriculum Development Model			x		
MBTI	x				
Processing Model				x	
VARK			x		x
Witkin's	x	x			

Most of the model above discusses individual preferences towards personality traits, individual abilities, learning activities preferences and learning approaches preferences; only VARK LS model uses sensory modality in their model, hence, most suited to RM.

2.5.3 The Importance of Learning Style in Reading Material Selection

Many research focused on providing various teaching strategies to accommodate students' different LS. They provided students with different type of teaching or learning strategies such as using power point, tutorial, group work, lectures, simulations, notes, games and virtual field trip. This may help students in the classroom, but not outside the classroom.

Students need extra RM for learning and they need to find the extra RM by themselves. Usually, students search RM by topic. As long as the content of the RM is related to the topic that they are looking for then it is enough for them. However, if the information obtained is not aligned to their LS, the students may find it difficult to understand the RM.

RMs can be presented in different ways. It depends on the authors. They organize and present the information according to how they like the information to appear. There are some authors who like to present information in textual form with a lot of examples, whilst others prefer to include graphics in appropriate places. However, it is difficult for students to learn when the RM is not presented in their preferred way (Gregorc, 1985). On the other hand, if the way information is presented in the RM matches with their LS, it enhances their learning capabilities.

However, most students are unaware of their own LS preferences (Rogers, 2009). Some of them know about LS, but do not use and take advantage of it in their learning process especially when it comes to choosing RMs. This leads to the mismatches between the LS of students and the presentation of the information in RM. It is necessary to make students aware of their own LS since most of them are unaware of what works for them and therefore are not realizing their potential because they are not studying in a way that best matches their LS (Honey & Mumford, 1992; Fleming, 1995).

Differences between students' LS and the way information is presented affects students' understanding of the subject and their attitude in class (Felder, 1995). Students may lose their focus, become demotivated, and assume that the course is too difficult for them, thus giving up on the course (Godleski, 1984; Renzulli, 1986; Felder & Silverman, 1988).

2.6 Information Seeking Tool

Information retrieval tool or Information seeking tool (IST) is an important element in information seeking. It is not just a system that stores and retrieves information. It also contains a set of components which are interrelated together to facilitate searching process. The components are document, indexing, vocabulary, searching, user-system interface, and matching components (Lancaster, 1979).

In this research, the term IST is used instead of an information retrieval tool, because this research only covers a few components of information retrieval tool such as document, user-system interface and mapping components. Other information retrieval components such as indexing, searching and vocabulary components are not covered.

Chowdhury (1999) described the basic process of IST as shown in Figure 2.5. A file of potential search term is produced from organized information. When a user submits a query, a comparison is made between the file and the query. A set of documents are then retrieved.

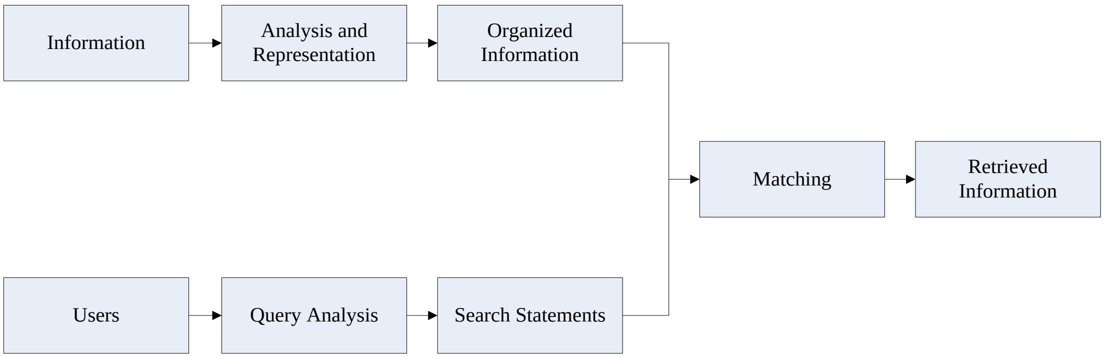


Figure 2.5: Basic Process of IST (Chowdhury, 1999)

ISTs differ in structure, in the way they function and utilize different methods and techniques for storing and retrieving information. A few examples of ISTs are described below.

- **Online public access catalogues (OPAC)**

  OPAC is a computerized catalogue containing bibliographic records of items in a library (Ariyapala, 2002). Students usually used OPAC to find books from library online before borrowing it. By using OPAC, students can access bibliographic records of all type of materials in the library. They can search more than one collection within the same library or in different libraries. OPAC provide a simple search interface that enable users to browse and search the collections. Search facilities provided by OPAC are browse and search, keyword and phrase search, Boolean and proximity searching (limited to keyword search option), subject heading such as Library of Congress Subject Headings searching, and searching records through selected keys such as author, title, International Standard Book Number or call number (Chowdhury & Chowdhury, 2003). Pendeta WebPAC as shown in Figure 2.6 is an example of OPAC from the University of Malaya Library.

Figure 2.6: Pendeta WebPAC Search Page

- **Internet search engine**

Internet search engine consists of a web page, images and files. It uses web crawlers or spiders to automatically retrieve information from millions of web pages on the web and then indexes the information. This makes Internet search engine the most comprehensive coverage of the web. An example of an Internet search engine is Google, shown in Figure 2.7. Each Internet search engine has its own specific set of retrieval features. Search facilities provided by most Internet search engines are word, phrase and natural language search options, special options for image, audio and video search, multilingual search and common search facilities such as Boolean search. They can restrict and rank searches with different criteria. Although Internet search engines provide the most innovative and the latest retrieval techniques, there are several issues regarding the quality of the information. Search engines index so many web pages automatically, hence they do not discriminate the qualities of the materials that are indexed (Cooke, 2001).

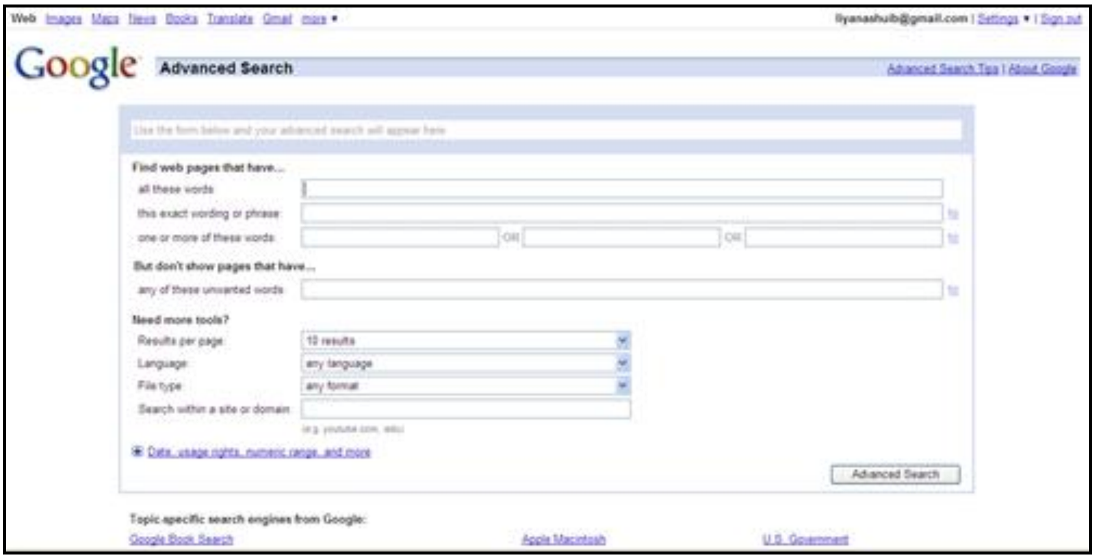


Figure 2.7: Google Search Page

- **Subject directory**

Subject directories are created manually by assigning the submitted sites to a suitable subject category by the directory developers. Search facilities provided by most subject directories are browse and search options, word, phrase and subject search options and ability to rank result by relevance option. Examples of subject directory are Yahoo and Britainnica.

- **Online database**

Online database are collections of scholarly journals that can be accessed online. It provides access to remote database through a database vendor or service provider. Format use is the same as the traditional printed journals and can be in the form of HTML, PDF or both. Chowdhury and Chowdhury (2003) noted that online databases fall into two major categories. One with printed counterparts like Journal of Documentation and the other is only available in electronic format such as D-Lib Magazine. Search facilities of online databases are browsing and search facilities, keyword and phrase search facilities, and common search facilities such as Boolean search, truncation, field search, limiting search and range search. It can be browsed by issue or the entire collection. Examples of online databases are Elsevier, IEEE and ACM.

- **Digital library**

Arms (2000) defined digital library as a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network. It provides a high quality resource within a particular subject area that is of interest to a specified audience (Cooke, 2001). It is filtered by library professionals and subject experts and added manually. Search facilities provided by most digital libraries are browse and search options, word, phrase and subject search options and common search facilities such as Boolean search. Digital library search functions are limited because it is designed to focus more on content. Examples of digital library are Dspace@UM and DigiLibraries.com.

2.6.1 Limitation of Existing Information Seeking Tools

Table 2.3 - 2.5 illustrates the differences between several IST in terms of the capabilities and functionalities provided. The comparisons are divided into three categories:

- Access

Denotes how the information is presented to the users. It could be in the form of full texts, as a list of bibliography or in the form of links to other websites or other databases.

- Search Strategies

Describes the techniques employed by the search engines to aid users in narrowing down their query or search results.

- Search Facilities

The number of different search categories available for the users to perform.

Table 2.3 shows the type of information that can be accessed by users. Internet search engine, online database and digital library can access document in the form of full text and metadata. OPAC can only access bibliographies while subject directories can only access links.

Table 2.3: Comparison of ISTs Access

	OPAC	**Internet Search Engine**	**Subject Directory**	**Online Database**	**Digital Library**
Full Text		x		x	x
Bibliography/ Metadata	x	x		x	x
Link			x		

Table 2.4 shows types of search technique adopted in ISTs. Internet search engine and online database are at the top of the list by providing the most search techniques. Subject directories and digital library only employ Boolean search and limiting search.

Table 2.4: Comparison of ISTs Search Strategies

	OPAC	**Internet Search Engine**	**Subject Directory**	**Online Database**	**Digital Library**
Boolean search	x	x	x	x	x
Proximity searching	x	x		x	
Truncation,		x		x	
Field search		x		x	
Limiting search	x	x	x	x	x

Table 2.5 shows the type of search facilities provided in IST.

Table 2.5: Comparison of ISTs Search Facilities

	OPAC	**Internet Search Engine**	**Subject Directory**	**Online Database**	**Digital Library**
Browse and search	x	x	x		x
Keyword and phrase	x	x		x	x
Subject headings	x		x		x
Natural language		x			
Retrieval Mode: Novice/Expert		x		x	
Ranked search result		x	x	x	x
LS					

From Table 2.5, we note that none of the existing ISTs has a function that matches RMs to students' LSs.

2.6.2 Personalized Information Seeking Tool

More specific ISTs are then developed to cater for individual LS. They are usually developed for education purposes. In this case, ISTs component is embedded in educational system such as e-learning. E-learning is an online delivery of information for purposes of education, training, or knowledge management (Sun & Xie, 2009).

In personalized IST, LS is used as one of the attributes in the retrieval process to accommodate student LS with learning material provided in the system (Bachari et al., 2011; Franzoni & Assar, 2009; Yang & Wu, 2009). By taking into account of students' LS, achievement in personalized IST is improved (Zapalska & Brozik, 2006). It eliminates *one size fits all* problem that gives the same learning materials to each student.

It is important to provide a personalized IST that can provide learning material for the varied needs of different students because one type of learning material does not fit all students. Many students have problems in grasping the contents and concept which leads to students’ confusion (Yang & Wu, 2009). Students learn more effectively when the instruction is matched with their individual LSs (Rasmussen, 1998). Several studies have revealed that learning can be enhanced through the presentation of materials that are consistent with a student’s particular LS (Budhu, 2002; Pen˜a et al., 2002; Stash et al., 2004).

There are many research that implemented LS in their personalized IST. Personalized IST facilitate students to learn better by using different way according to students’ preference (Bachari et al., 2011). In these personalized IST, the learning materials are then presented in the way that best fit the LS of students. Previous studies have reported that by using different learning material, student learning abilities can be enhanced (Franzoni et al., 2008; Yang & Wu, 2009).

In recent years, there has been an increasing amount of personalized IST developed by researchers. Table 2.6 shows personalized IST from year 2008 to 2011. These personalized IST used various LS models to match students with learning material. Most of them used well known LS model such as Kolb (Yang & Wu, 2009), VARK (Apriyani & Hasibuan, 2008; Hassan et al., 2010), MBTI (Bachari et al., 2011) and FSLSM (Franzoni & Assar, 2009; Klašnja-Milićević et al., 2011; Savic & Konjovic, 2009; Schiaffino et al., 2008). Only Cheng (2009) used a combination of multiple LS preference in his study. Bachari et al. (2011), Franzoni & Assar (2009), Klašnja-Milićević et al. (2011) and Schiaffino et al. (2008) used formative assessment such as short quiz, assignments and homework as their learning resources. Whilst others used learning objects such as video (audio/animation), text (RM, power point) and graphic (chart, image, diagram). Formative assessment and learning object used in the personalized e-learning above are specific to only one LS preference.

Most of the personalized IST proposed learning strategy based on learning objects to student. Students are routed to specific learning strategies based on their LS preference. Savic & Konjovic (2009) and Schiaffino et al. (2008) proposed learning path that recommend a route to follow learning content. Klašnja-Milićević et al. (2011) used both in his system, Protus.

Table 2.6: Personalized IST

Author	LS model	Learning Resources	System Output
Bachari et al. (2011)	MBTI	Formative assessment	Learning strategy
Klasnja-Milicevic et al. (2011)	FSLSM	Formative assessment	Learning path and learning strategy
Hassan et al. (2010)	VARK	Learning object	Learning strategy
Yang & Wu (2009)	Kolb	Learning object	Learning strategy
Savic & Konjovic (2009)	FSLSM	Learning object	Learning Path
Cheng (2009)	Multiple style	Learning object	Learning strategy
Franzoni & Assar (2009)	FSLSM	Formative assessment	Learning strategy
Apriyani & Hasibuan (2008)	VARK	Learning object	Learning strategy
Schiaffino et al. (2008)	FSLSM	Formative assessment	Learning path

All the previously mentioned personalized IST has some drawbacks. One of the problems was the system only used LS to accommodate students with learning material that is specially designed for the system. Researchers have to develop various learning objects of the same subject to meet the need of students' different LS (Bachari et al., 2011). This task is complicated and increases cost and time.

2.7 Theoretical Framework

This research was based on two theoretical areas, namely the Principle of Least Effort (Zipf, 1949) in information seeking and the theory of Multiple Intelligence (Gardner, 1993) in learning style. Both were discussed in Sections 2.3 and 2.5, respectively. Marchionini's Information Seeking Model (Marchionini, 1995) was used to adapt the Principle of Least Effort. This model is used in IST tools and involves several steps during the information seeking process. In this research, we only covered Documents, User-system interface, and Matching components of the IST tool.

The Theory of Multiple Intelligence (Gardner, 1993) is based on the belief that each individual has a unique intelligence or preference. The VARK LS model was derived from this theory and used as the LS model in this research. The theoretical framework of this research is built on these theory and model as shown in Figure 2.8.

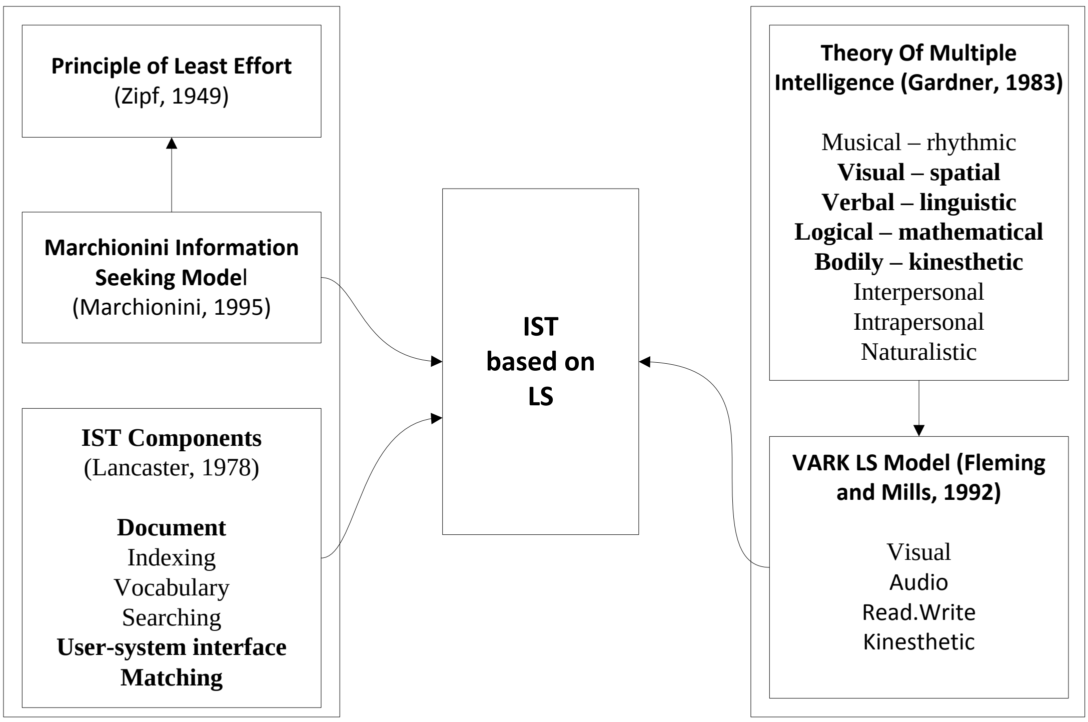


Figure 2.8: Theoretical Framework

2.8 Summary

Based on the literature review, research gaps and problem statements were derived. Figure 2.9 shows how the components of the literature review, research problem statements, questions and gaps, relate to each other.

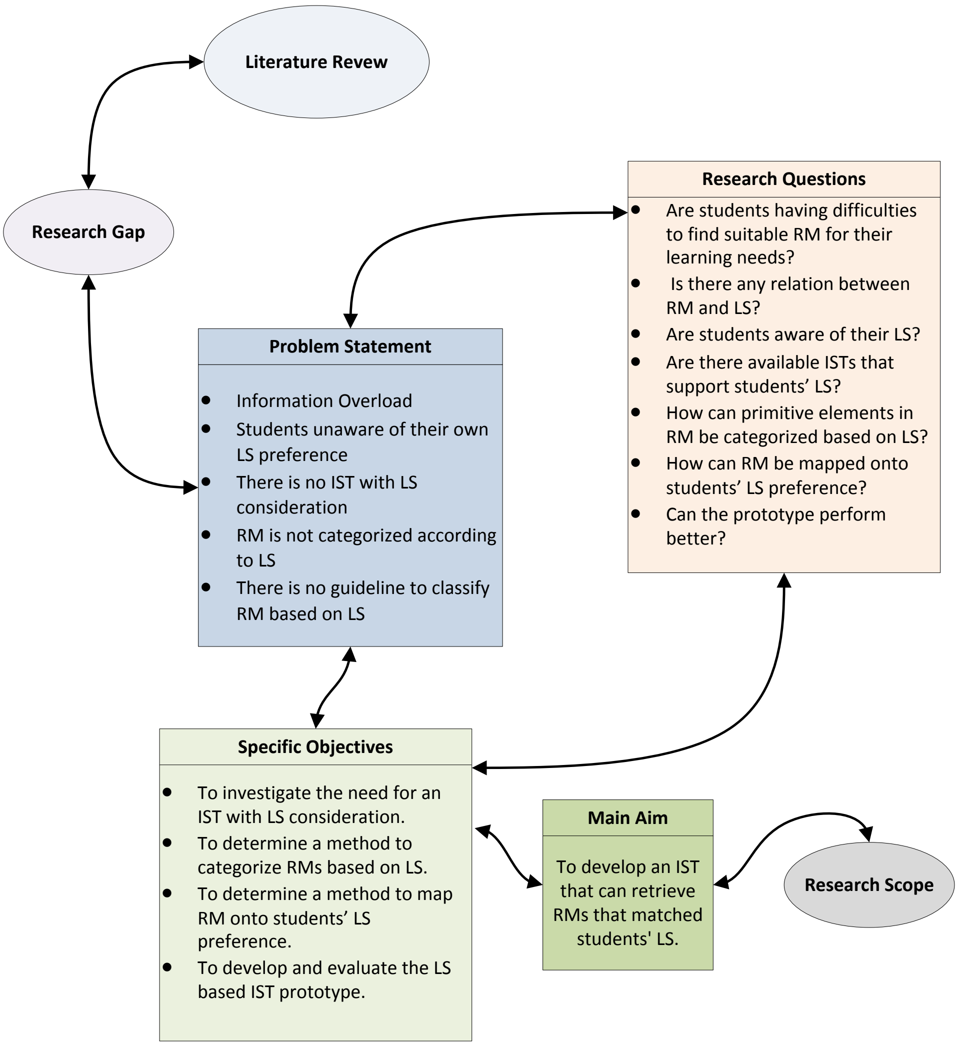


Figure 2.9: Relationships between problem statements, objectives questions and gaps

The findings of this research suggest that students have difficulties in finding suitable RMs due to:

- Information overload
- Students being unaware of their own LS preference
- No IST with LS consideration
- RMs not being categorized according to LS
- No guideline to classify RM based on LS

It is therefore important to take into account students' LS when developing an IST. Currently, there are no tools available that can help students choose RMs based on their LS. Available ISTs do not include LS as one of the attributes during the retrieval process. Several personalized ISTs, such as e-learning, include students' LS components. However, they only focus on providing different learning strategies or learning materials, instead of finding suitable RMs to accommodate students' LS. The following chapter discusses the steps taken during this research.

CHAPTER 3

RESEARCH DESIGN

3.1 Introduction

Chapter 3 describes the research design used in this study. A research design is a collection of all of the techniques, methods, and procedures used to carry out our research, and justify the process used to examine the research's objectives. Research objectives for this research are:

1. To investigate the need for an IST with LS consideration
2. To determine a method to categorize RMs based on LS
3. To determine a method to map RM onto students' LS preference
4. To develop and evaluate the LS based IST prototype

This research employs quantitative methods that aim to answer the following research questions:

- Are students having difficulties to find suitable RM for their learning needs?
- Is there any relation between RM and LS?
- Are students aware of their LS?
- Are there available ISTs that support students' LS?
- How can primitive elements in RM be categorized based on LS?
- How can RM be mapped onto students' LS preference?
- Can the prototype perform better?

The research design of this research consists of the following three stages:

- **Identification of the problem**

Identification of the problem involves a literature review and a preliminary study. A literature review was conducted to develop an understanding of previous works related to information seeking, IST, document analysis, the concept of learning styles, and their relationship within the student's information seeking process. Next, a preliminary study was carried out to confirm the findings obtained from the literature review.

- **Design and development of LSIST**

Based on the findings from the first stage, architecture for the LSIST was designed. This system was developed using SSADM methodology. This will be discussed in detail in Section 3.3.

- **Evaluation of LSIST**

The prototype will be evaluated by the user to see how this prototype can meet the objectives of this research. This prototype evaluation will be achieved through experiments and surveys, and will be discussed in detail in Section 3.4.

3.2 Identification of the Problem

The first stage of this research is the identification of the problem. This involves two processes, namely literature review and a preliminary study.

3.2.1 Literature Review

In the literature review, the requirement of the research will be defined. This process defines the problem and collects the data needed for the process of solving the problem. A literature review was used to study the domains involved in this research. The domains involved are illustrated in the mind map below (as discussed earlier in Chapter 2).

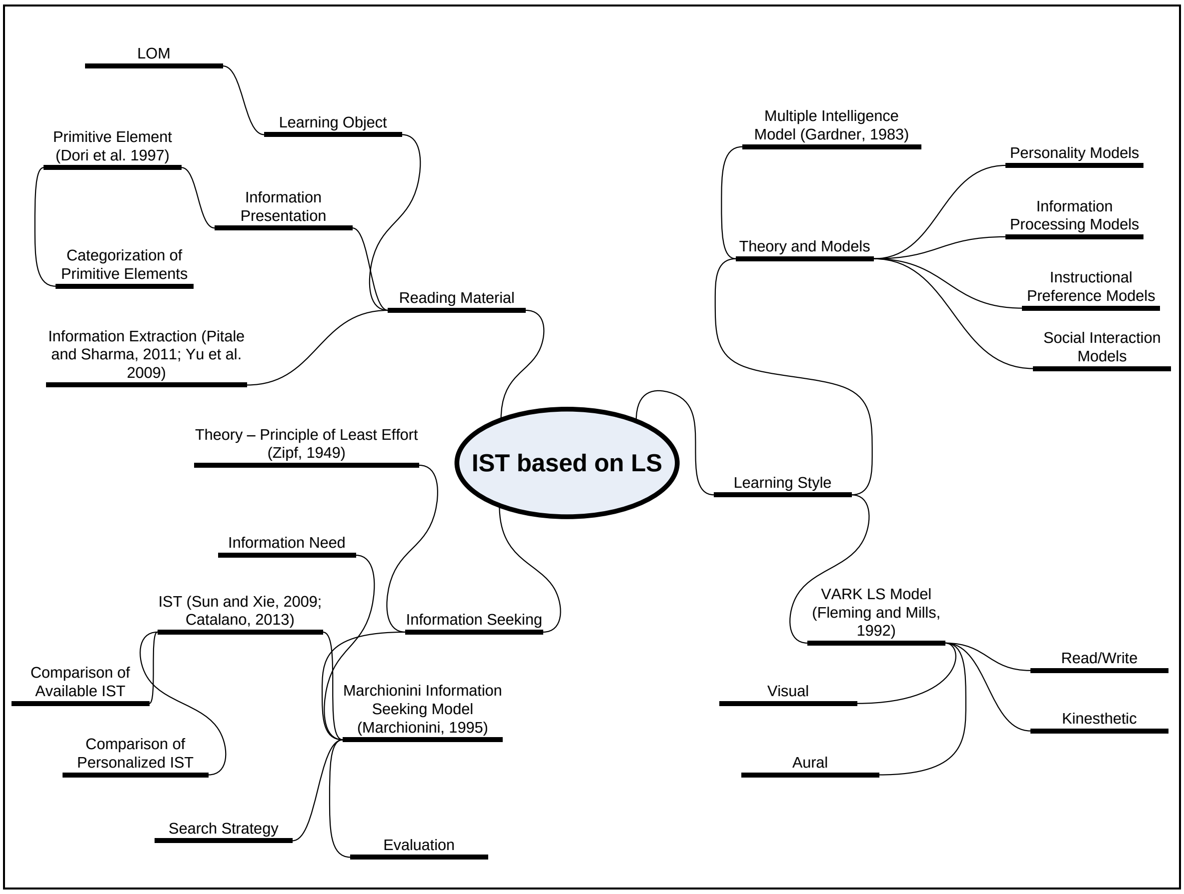


Figure 3.1: Mind map of literature review

A theoretical framework for this research was developed based on previous literatures (as shown in Section 2.8).

3.2.2 Preliminary Study

In order to confirm the findings of students' difficulties in finding suitable RM from the literatures, we conducted two preliminary studies. We chose surveys, since surveys are widely acknowledged as an acceptable research tool to gain a reflection of the attitudes, preferences, and opinions of the community (Rea & Parker, 2005). The main instrument used in the design of this survey was questionnaire. Questionnaire is chosen because:

- It facilitates the collection of large amount of data with less effort in a short period without sacrificing the efficiency and accuracy.
- It provides an entirely standardized measuring instrument because the questions are phrased exactly in the same way for all respondents (Sapsford, 2007).
- The respondents are postgraduate students. With the experiences in research, most of them are expected to have answered some sort of questionnaires before and should not have any difficulties in answering the instrument distributed.
- The respondents have greater confidence in their anonymity and thus feel free to express their views.

3.2.2.1 Questionnaire 1: A Survey on Students' Information Seeking Behaviour

Various information seeking models from the literature were considered before designing the questionnaire. The information seeking model by Marchionini (1995) was chosen, because its information seeking process in electronic environments was closely related to the needs of this study.

The objectives of this survey are:

- To study the information seeking behaviour of Computer Science students.
- To ascertain the problems faced by students in seeking information.

The process involved for this questionnaire is discussed below:

- **Population and Sampling**

The unit of analysis was individual postgraduate students. With a total of 498 graduate students currently active in FCSIT, a representative sample size of 140 was deemed reasonable to give a satisfactory response rate of 28.1 per cent. For this study, convenience sampling was used. In convenience sampling, a sample is selected from elements of a population that are easily accessible and willing to be studied (Creswell, 2009). The respondents provided a reasonable representative profile of all postgraduate students. Replies were obtained from various age groups, genders, types of study, modes of study, and nationalities.

- **Instrument Development**

This questionnaire was a four-page self-administered questionnaire. The questionnaire was divided into the following four parts:

Part 1: Demographic details, such as gender, age, type of study, mode of study, and nationality (six questions)

Part 2: Problems respondents experienced during the information seeking process (four questions)

Part 3: How respondents searched for information (four questions)

Part 4: How respondents evaluated and verified the information that they had retrieved (five questions)

Table 3.1: Questionnaire 1 Items

Part	Item on the Questionnaire	Type of scale	Remarks
1	1,2,3,4,5,6	Categorical	Demographic details
2	7, 9, 10	Categorical	To ascertain the problems faced by students in seeking information
	8	Continuous	
3	11	Continuous	To study the information seeking behavior. The questions are constructed using Marchioninis' Information Seeking Model (1995)
	12, 13, 14	Categorical	
4	15, 16, 17, 18	Categorical	
	19	Text	Comments

- **Pre-testing**

The initial questionnaire was pre-tested using Cooper and Schindler's (2006) pretesting method; wherein the questionnaire was reviewed by two senior researchers from the same field, in order to ensure that the questions were both valid and accurate. After the questionnaire was revised, a pilot-test was performed using a small sample of the respondents (n=10), using computer science postgraduate students from other universities.

- **Data Collection**

Full-scale data was collected, using an online questionnaire that was distributed to all postgraduate students via an email attachment. Two weeks after the initial email was sent, a follow-up email was sent to remind respondents to complete their questionnaire. A second follow-up email was sent to those respondents who did not respond within one month. Note: no incentives were provided to respondents to complete the questionnaire.

- **Data Analysis**

The information obtained from the returned questionnaires was coded and transferred into the Statistical Package for Social Science (SPSS) version 17 for Windows. The questionnaire is appended in Appendix A. Each respondent is labeled as S1, S2 and so on in SPSS. Data analysis, such as descriptive statistics and Chi square testing, was used to obtain the results.

3.2.2.2 Questionnaire 2: A Survey on Students' Learning Style Awareness

This questionnaire used the VARK LS test to identify students' preferences (Fleming, 2010). The objectives of the questionnaire are:

- To investigate students' awareness of LS,
- To identify students' LS preferences.

The process involved for this questionnaire is discussed below:

- **Population and Sampling**

Data samples were gathered from postgraduate students in the Faculty of Computer Science and Information Technology (FSCIT), University of Malaya. For this study, convenience sampling was used, where the sample is selected from the elements of a population that are easily accessible and willing to be studied (Creswell, 2009). The unit of analysis was individual postgraduate student. With a FSCIT postgraduate students' population of 498 students, a representative study population of 111 students was used, giving a response rate of 23 per cent which provided a reasonable representative profile of all postgraduate students from various levels of study and nationalities.

- **Instrument Development**

This questionnaire was a four-page self-administered questionnaire. The questionnaire was divided into the following three parts:

Part 1: Demographic details, such as gender, nationality, and level of study (three questions)

Part 2: Respondents' LS awareness and preference (five questions)

Part 3: VARK LS test (from the VARK website), consisting of 16 multiple choice questions with four answer selections, corresponding to the four sensory modalities (Fleming & Mills, 1992) (sixteen questions)

Table 3.2: Questionnaire 2 Items

Part	Item on Questionnaire	Type of scale	Remarks
1	1,2,3	Categorical	Demographic details
2	4, 5, 6. 7	Categorical	To investigate students awareness of their LS
	8	Text	
3	1-16	Categorical	VARK LS test (Fleming & Mills, 1992)

- **Pre-testing**

The initial questionnaire was pre-tested using Cooper and Schindler's (2006) pretesting method; wherein the questionnaire was reviewed by two senior researchers from the same field, in order to ensure that the questions were both valid and accurate. After the questionnaire was revised, a pilot-test was performed using a small sample of the respondents (n=10), using computer science postgraduate students from other universities.

- **Data Collection**

After pre-testing in March 2011, the final questionnaire was distributed to the intended respondents between April and July 2011 as an offline questionnaire.

- **Data Analysis**

The questionnaire is appended in Appendix B. Each respondent is labeled as S1, S2 and so on in SPSS. Data analysis, such as descriptive statistics and Paired Sample T-Tests, was used to obtain the results.

3.3 System Design and Development

This system was developed using the Structured Systems Analysis and Design Method (SSADM) (Weaver et al., 1998). SSADM is highly structured and well documented. It was chosen because it provides very detailed rules and guidelines to study existing systems. In this research, the available IST process was studied; from which a new process was included in the proposed prototype. The use of a data flow diagram, to show data flowing in and out of the system, helps researchers to design prototypes more effectively. This was another reason why this method was chosen.

SSADM contains five modules; which are feasibility study, requirement’s analysis, requirement’s specification, logical system specification, and physical design (as shown in Figure 3.2). Module function and deliverables are summarized in Table 3.3. The design and development will be discussed in detail in Chapter 6.

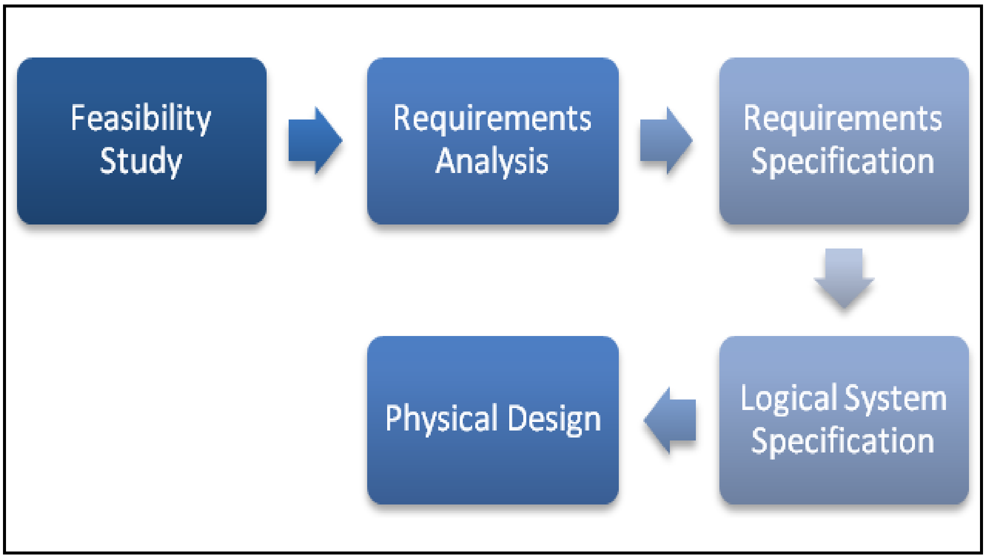


Figure 3.2: SSADM Phase

Table 3.3: SSADM Phases, Function and Deliverable (Ashworth & Slater, 1992)

Module	**Function**	**Deliverable**
Feasibility Study	To identify scope, problem and study current requirement	Problem from current system
Requirement Analysis	To define in overview the requirements for the new system by extracting the essential functionality and data from the current system and documenting the users' requirement for the new system	Current Services Description, Requirement Catalogue and User Catalogue
Requirement Specification	To analyze in detail the requirements for the new system	Required ERD and DFD, User Roles and User Roles/Function
Logical System Specification	To provide a detailed specification of the processing and dialogue requirements for the new system	Technical System Option and Requirement of Design
Physical Design	To act as a bridge between the logical design and the construction phase of a physical design	Database, Interface and Documentation.

3.3.1 Proposed Architecture

This stage starts by mapping the requirements delivered by the problem identification stage into the system's architecture. The architecture defines the components and the process involved (as shown in Figure 3.3). In order to develop an IST that incorporates LS into its searching process, the architecture of the proposed tool needs to provide a clear method for the design of the LSIST. Several components were added to the basic process of IST (as discussed in Section 2.6) as follows:

- Students' LS profile

To identify students' LS preference based on the LS test.

- RM Classification

To extract LS value in RM.

- Matching

To match RM, based on student' LS.

The proposed architecture involves two types of users, namely administrators and students. Administrators provide RM to the tool. RM, used as a source for this prototype, was restricted to those formatted in PDF. Figure 3.3 shows the proposed architecture.

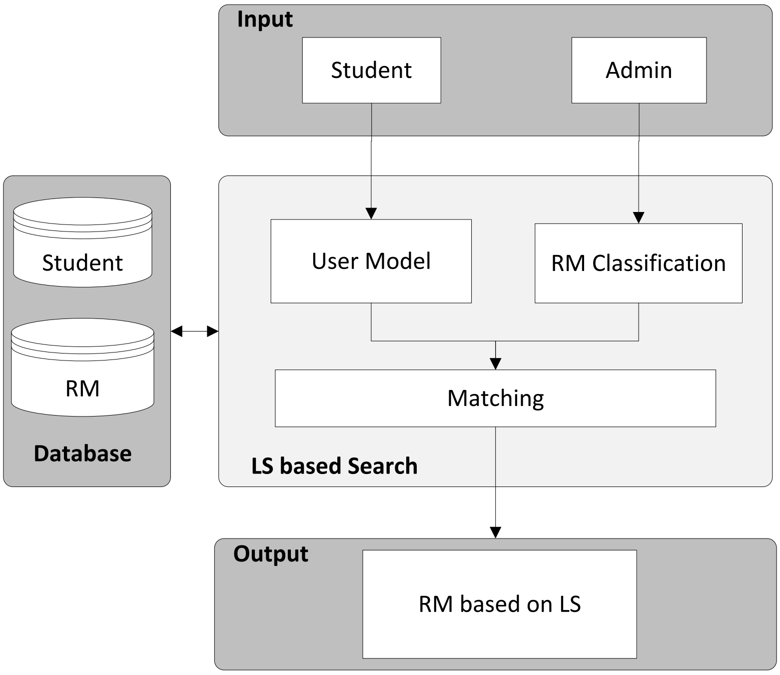


Figure 3.3: LSIST Architecture

LSIST architecture consists of the following four main modules; Input, LS based Search, Database, and Output:

- **Input Module**

The input module receives inputs from users, such as students' profile, LS preferences, search query, RM profile, and LS value information.

- **LS based Search Module**

The LS based Search retrieves RM based on keyword and LS. This module contains three main components, namely User Model, RM Classification, and Matching. The details of this module will be discussed later in Chapter 5.

- **Database Module**

This module comprises two databases, which are Student Database and RM Database

- **Output Module**

The recommended RM is displayed in the Output module.

3.4 Evaluation of LSIST

Experimental research was chosen as the evaluation method. It is a traditional approach to conduct a quantitative research. It is intended to establish possible causes and effects between the independent and dependent variables (Creswell, 2008). There are two types of experimental research, namely true experiment and quasi-experiment. True experiment uses the random assignment of subject to treatment conditions, while quasi-experiment uses a non-randomized design (Keppel, 1991). This research used quasi-experimental design to control all of the variables that influence the outcome; except for the independent variable.

3.4.1 Quasi-Experiment

In this research, Quasi-Experimental design i.e., Pre-test/Post-test, Non-equivalent Multiple Group Design, was used as the type of experimental design. It is a type of evaluation that aims to determine whether an intervention has the intended effect on a study's respondents. According to Wiersma and Jurs (2009), there are three major design types in Quasi-Experimental design, which are Non-equivalent Control Group Design, Time Series Design, and Single-Subject Design.

This research used Non-equivalent Control Group Design, because it involves the comparison of groups and not time series. Non-equivalent Control Group Design consists of two types of design; Post-test-Only and Pre-test/Post-test. This research used Pre-test/Post-test design to compare results between before and after intervention. Figure 3.4 shows the processes involved for the experiment design; the details of which are explained in Section 6.2.3.

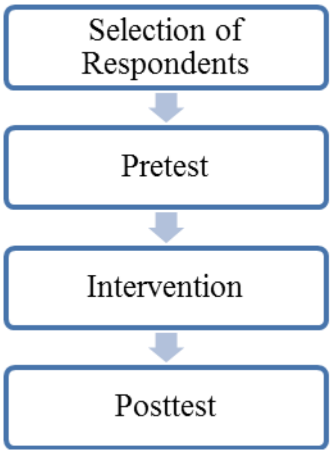


Figure 3.4: Pretest/Posttest, Non-equivalent Multiple Group Design

To conduct the experiment, respondents and materials were selected. The evaluation process procedure is discussed in Section 3.4.2.

- **Respondents**

Fifty Computer Science postgraduate students were selected to participate in this experiment. They were divided into five experimental groups, based on their LS preferences; as shown in Figure 3.5; i.e., each LS preference is represented by ten respondents.

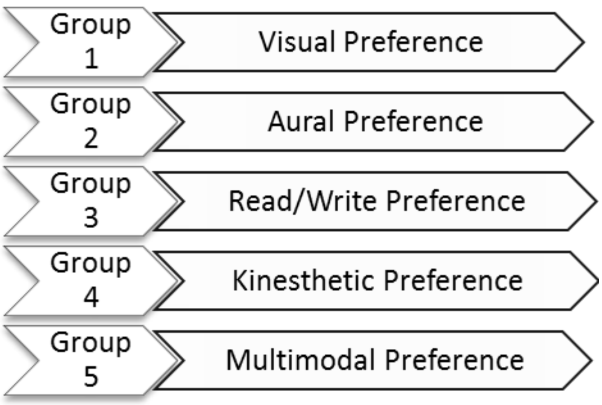


Figure 3.5: Experimental Groups

To select a respondent, a quota sampling technique was used. This technique allows the researcher to set the number of respondents needed in a sample. Next, convenience sampling was used to locate the participants to meet the quota.

- **Material**

RM was chosen as the material in this research, because it is an easily accessible learning material. It is also a reusable learning object and can be used and created without difficulty. For the evaluation, 77 RMs (in PDF), for various topics such as e-learning, data mining, artificial intelligence, research, system development, and knowledge management, were uploaded into the prototype. These RMs contained various primitive elements (as shown in Appendix E).

- **Procedure**

The design of this experiment was based on experimental groups. The design allows us to make inferences on the effect of the intervention by looking at the difference in the Pre-test and Post-test results (as shown in Figure 3.6). In this case, intervention is the identification of students' LS and their use of LS information in the searching process. Results from Pre-test and Post-test were compared using statistical analysis. Instructions for the prototype's evaluation are shown in Appendix F.

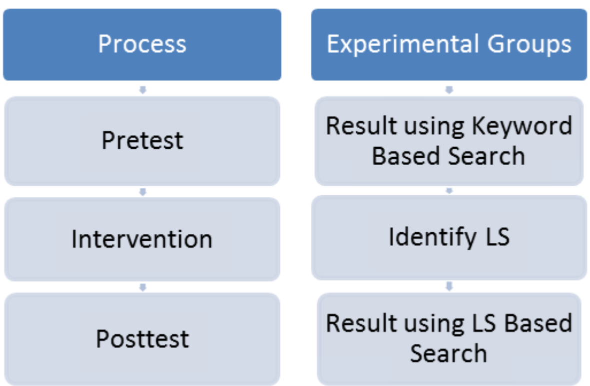


Figure 3.6: LSIST Evaluation Process

Respondents were asked to run the prototype. They went through the steps below:

Step 1: Search for RM using *Keyword based Search* (Pretest)

Step 2: Answer a LS test to identify their LS preference (Treatment)

Step 3: Search for RM using *LS based Search* (Posttest)

Step 4: Give feedback

After the RMs in step 1 and 3 are retrieved, the respondents view the RM and rate the RM according to their preference based on how information is presented in the RM. The results of the evaluation are discussed in Chapter 7.

3.4.2 Survey

A simple survey on the respondents' opinion about the prototype was conducted. It comprises the following five questions:

- Is this tool helpful in providing suitable reading material to you?
- Is there any difference between *LS based Search* and *Keyword based Search*?
- Does *LS based Search* provide better results?
- Do you like the way information is presented in the reading materials suggested by the *LS based Search*?
- Is the reading material retrieved from the *LS based Search* easier to understand than those given by *Keyword based Search*?

This survey provided feedback for the prototype. Students answered these questions online after finishing the experiment.

3.5 Summary

This chapter discussed the research design for this study. Chapter 2 and Chapter 4 discussed the findings from the Problem Identification stage. The Design and Development of the LSIST stage is presented in Chapter 5 and Chapter 6, respectively. Chapter 7 covers the final stage; i.e., the Evaluation of LSIST. Table 3.7 shows the research framework and summarizes the research outcomes and deliverables of each stage.

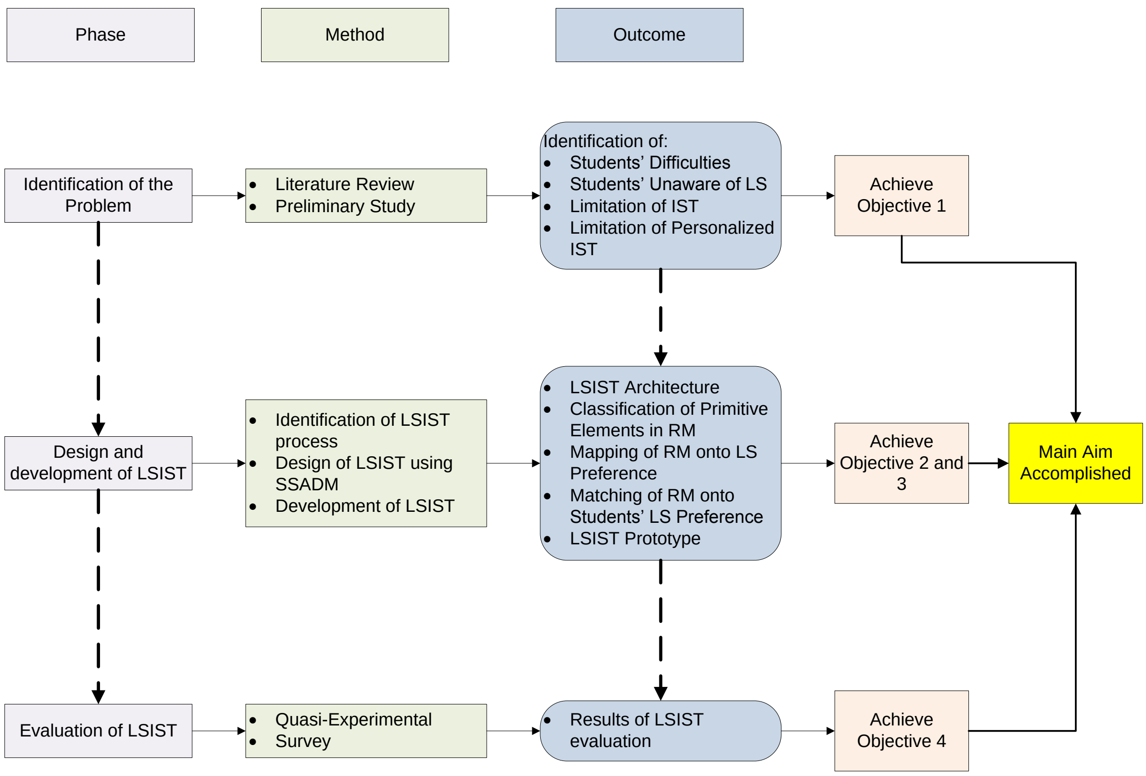


Figure 3.7: Research Framework

CHAPTER 4

STUDENTS' INFORMATION SEEKING BEHAVIOUR AND THEIR LEARNING STYLE AWARENESS

In Chapter 2, we discussed student difficulties in information seeking, LS and ISTs. To confirm the findings of students' difficulties in finding suitable RM from the literatures, we conducted two surveys. Two sets of questionnaires were developed and administered. For each set, a structured self-complete questionnaire was used.

4.1 Questionnaire 1: A Survey on Students' Information Seeking Behaviour

Questionnaire 1 is conducted to study the information seeking behaviour of Computer Science students and to ascertain the problems faced by students in seeking information. The details of the survey design is discussed in Section 3.2.2.1.

4.1.1 Results

A total of 140 respondents responded to the questionnaire. 55.7% of respondents are Male. The respondents are categorized into the 25 years or below (20.7%), 26-30 years (44.3%), 31-40 years (22.9%) and 41 years or above (12.1%). The modes of study are Masters by Coursework (56.4%), Master by Research (34.3%), and PhD (9.3%). The majority of respondents are fulltime students (84.3%), 52.1% are Malaysians and 47.9% are foreign students from Iraq, Iran, Indonesia, Bangladesh and Sri Lanka. Table 4.1 shows the distribution of the respondents' demographic profile.

Table 4.1: Respondents' Demographic Profile

Demographic Profile (n=140)		Responses N (%)
Gender	Male	78 (55.7%)
	Female	62 (44.3%)
Age	25 years or below	29 (20.7%)
	26-30 years	62 (44.3%)
	31-40 years	32 (22.9%)
	41 years or above	17 (12.1%)
Types of study	Masters by Coursework	79 (56.4%)
	Master by Research	48 (34.3%)
	PhD	13 (93%)
Mode of study	Full time student	118 (84.3%)
	Part time student	22 (15.7%)
Nationality	Malaysian	73 (52.1%)
	Non-Malaysian	67 (47.9%)

a. Students' Information Seeking Behaviour

Marchionini (1995) proposed that information seeking is a special case of problem solving and it consists of the steps illustrated in Figure 2.1 and further explained below.

1. Recognize and Define Problems

This question is meant to determine whether respondents can recognize or define their problem while seeking for information. Results in Figure 4.1 show that most of the respondents reported that they do not have any difficulties in recognizing and defining their problem when seeking for information.

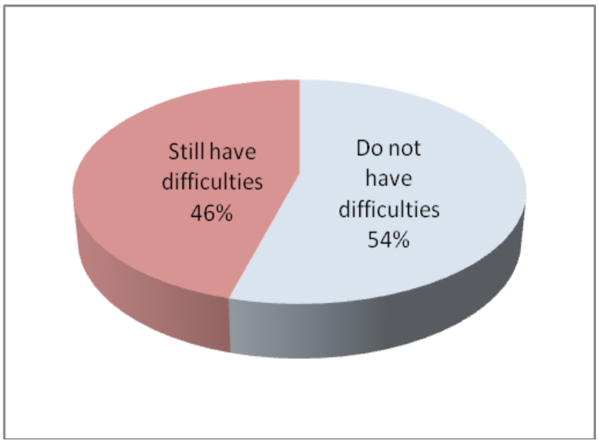


Figure 4.1: Having Difficulties in Defining the Problem

2. Selection of ISTs

There are various types of ISTs available for respondents to use to retrieve desired information. First, respondents were asked whether they use a variety of ISTs or only one type of IST to seek for information. The findings revealed that the majority (56.4%) of the respondents had *Very Often* used more than one IST as shown in

Figure 4.2. There were only two respondents who rarely use more than one IST. Both of them are from the Masters by Coursework program.

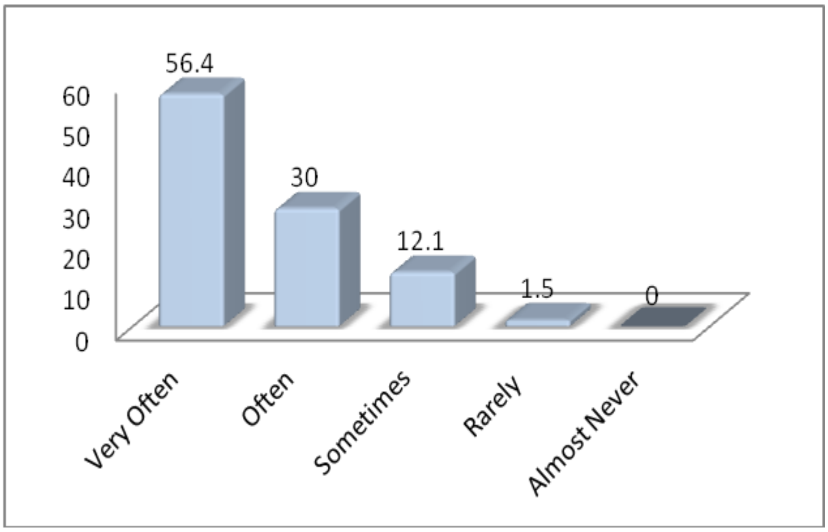


Figure 4.2: Use a Variety of ISTs

Table 4.2 demonstrates that Internet search engine is the most popular IST with 90% of respondents using it as their first choice. It is then followed by digital library and online databases with 67.8% of respondents reporting as using them frequently. OPAC was found to have the lowest usage for first choice with only 26.4%.

Table 4.2: First Choice of IST Used

	First Choice of IST Used (%)			
	Internet Search Engine	Digital Library	Online Database	OPAC
Frequently	90	67.8	67.8	26.4
Occasionally	8.6	22.1	22.1	41.4
Rarely	1.4	10.1	10.1	32.2
Total	100	100	100	100

This study reveals that most respondents use several ISTs even though most of them demonstrated that they use the Internet search engine as their first method, but they only use it as initial step to learn about the subject matter. After using the Internet search engine, they use other IST to find more information especially the trustworthy ones like digital library and online databases. This behaviour is supported by the following respondents' comments:

- *"There are many resources available now and it is difficult to read all that is being published. One need to be an expert to know how to select the most relevant resources based on the study area and not gets carried away by so much information. Better to begin with Google Scholar and then get the articles from library's subscribed databases or via document delivery." (S17)*
- *"Normally, I use Google and the university main library to get the information that I seek" (S14)*
- *"Mostly I use internet search engine (Google Scholar)."(S36)*

3. Query Formulation

In order to search information effectively, students must be able to identify and understand their problem and research area. They should also be able to formulate a query that is related to their problem area. From the findings, 77.1% of respondents are *Often* able to formulate their keywords with another 80.7% indicating the *Often* use of combined keywords. Only a small percentage (19.3%) of the respondents reported that they *Rarely* formulate an appropriate keyword using a combination of keywords as shown in Table 4.3.

Table 4.3: Respondents Ability to Formulate Keywords

Respondents ability to formulate appropriate keywords	**Responses N (%)**
Often	108 (77.1%)
Rarely	32 (22.9%)
Respondents ability to combine appropriate keyword	**Responses N (%)**
Often	113 (80.7%)
Rarely	27 (19.3%)

4. Duration to Find Information

Nearly all of the respondents (94.5%) reported that information seeking consumes more time than they expected. Detailed result is shown in

Figure 4.3.

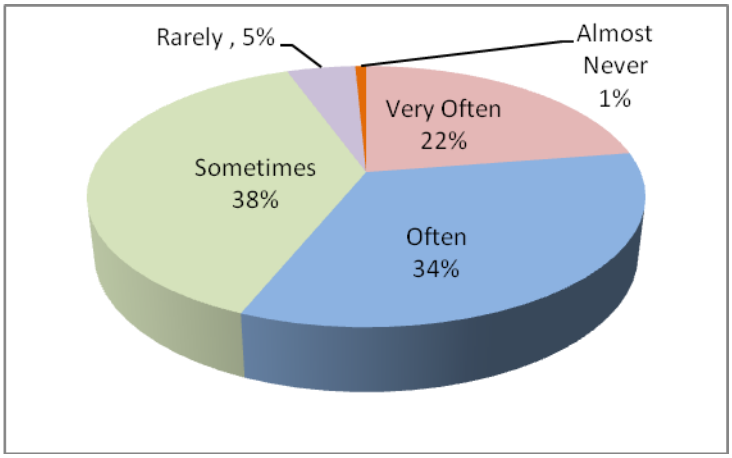


Figure 4.3: Information Seeking Consumes More Time Than They Expected

Figure 4.4 displays the duration respondents take to search for information. About 47.3% of the respondents took one week to find information while 41.1% took a day. 7.8% of the respondents spent within a month and 2.3% spent over a month to find information. This finding reveals that respondents spent a reasonable time to find information although some commented that they required more time to filter and evaluate the information.

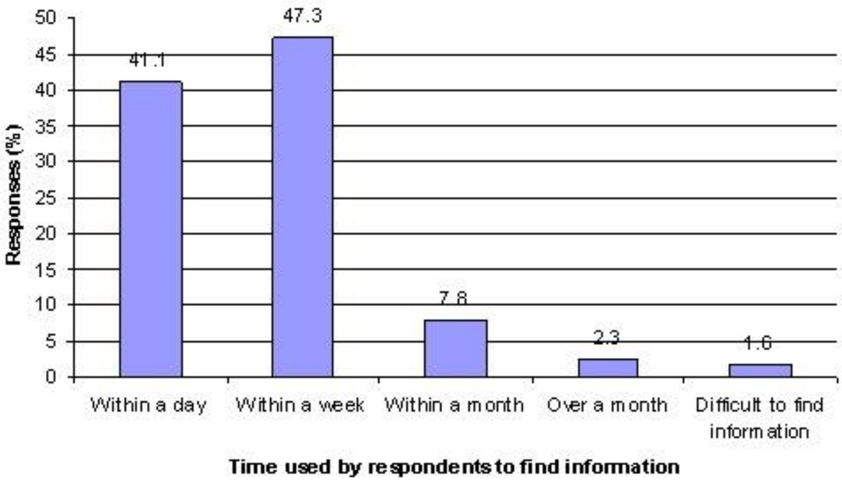


Figure 4.4: Duration Taken to Find Information

5. *Satisfaction with the Information Found*

Respondents' satisfaction with the information found is measured by evaluating if the information found matches their preferences. 70% of respondents say that they *Frequently* feel satisfied with the information they retrieved. The respondents then were asked if they could complete their research with the information they found. The responses were mostly positive where more than two thirds (77.8%) of the respondents indicate *Frequently* while the remaining 22.2% indicated *Rarely*.

In addition, respondents were asked the actions they would take if they cannot find the information at the first attempt. The majority opted to try another IST (85.7%), try another combination of keyword (80.7%), consult an expert (58.6%) and discuss with friends (50.7%). These are shown in Table 4.4. For this question, the respondents were allowed to select more than one option.

Table 4.4: Actions Taken By Respondents

Actions Taken By Respondents (N=140)	**Responses N (%)**
Try another information resources	120 (85.7%)
Try another combination of keyword	113 (80.7%)
Consult an expert	82 (58.6%)
Discuss with friends	71 (50.7%)
No Action	0

b. Problems in Information Seeking

The study investigates the problems faced by respondents while seeking for information. The result shows that respondents *Occasionally* (53.5%), *Frequently* (29.5%), *Rarely* (14.7%) and *Never* (1.6%) experienced problem. Table 4.5 ranks the problems faced by the respondents while seeking information. The findings revealed that 73.6% of the respondents selected *It is difficult to deal with the large amount of information available* as the main problem.

Table 4.5: Problem Faced by Respondents

No	**Problems (N=140)**	**Responses N (%)**
1	It is difficult to deal with the large amount of information available	103 (73.6%)
2	It is difficult to ensure that the information sources are trustworthy	92 (65.7%)
3	It is difficult to understand the information found	77 (55%)
4	It is difficult to know where to find relevant information	57 (40.7%)
5	It is difficult to categorize my information needs	63 (45%)
6	It is difficult to know how to access the information sources	48 (34.3%)
7	It is difficult to find the information that is relevant to my search subject	46 (32.9%)

Note: Respondents could select more than one option

Chi-square tests were performed to examine the relation between demographic profiles and problems in information seeking.

- **Relation between problems and gender**

  There is no significant relationship between problems and gender, χ^2 (4, N = 129) = 5.32, p = 0.26.

- **Relation between problems and age**

  There is no significant relationship between problems and age, χ^2 (12, N = 129) = 8.71, p = 0.73.

- **Relation between problems and types of study**

  There is no significant relationship between problems and types of study, χ^2 (12, N = 129) = 9.24, p = 0.68.

- **Relation between problems and mode of study**

  There is no significant relationship between problems and mode of study, χ^2 (4, N = 129) = 4.18, p = 0.38.

- **Relation between problems and nationality**

  There is no significant relationship between problems and nationality, χ^2 (4, N = 129) = 3.32, p = 0.51.

The results show that there is no significant relationship between demographic profile and problems.

4.1.2 Discussion

The first objective of this survey is to study computer science students' information seeking behaviour. This involves investigating:

- Their skills in recognizing and defining their problems
- The types of IST they normally used
- Their ability to formulate a query that is related to their problem
- The duration they need to find information
- Satisfaction with the information found

The findings show that some of them can recognize and define their problems but others still have difficulties at this initial stage. This depends on how they understand their information needs. The findings demonstrate that most of the participants use several ISTs. This study compares four ISTs; OPAC, Internet search engine, online database and digital library. Based on the study, it can be concluded that postgraduates still rely on Internet search engine to get information for their academic work, despite their level of knowledge and experience in seeking information. This findings are consistent with the findings of Becker (2003), Barrett (2005), Saiti & Prokopiadou (2008) and Julien & Barker (2009). They revealed that the Internet search engine especially Google is the most preferred IST.

While it is anticipated that postgraduate students use online databases to seek for journal articles for their research work, this is different with the findings from undergraduate studies where they were found to prefer books which are more relevant to their needs as stated by Barret (2005) and Saad & Zainab (2009). This study revealed that most participants use a variety of ISTs even though most of them demonstrated that they use the Internet search engine as their first source, but they only use it as the first step to know about the subject matter. Once they are acquainted with the subject matter, they use other more reliable sources to gain more understanding, by perusing digital library and online database. For example, postgraduate students of computer science usually start with Internet search engine such as Google to find general information and then turn to online database and digital libraries to find scholarly materials that are more trustworthy for their research. See Figure 4.5.

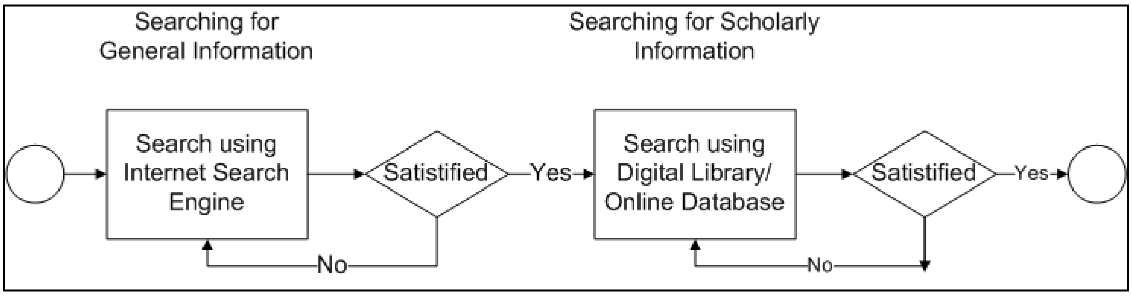


Figure 4.5: Respondents' Information Seeking Behaviour

The students do not face difficulties in finding and combining keywords to find their information. They are also able to conduct an information seeking session in a short period. If they encounter any problem in seeking information, they did not give up and repeated the search further until they are satisfied with their results.

The second objective is to investigate the problems faced in information seeking process. The data collected and analyzed showed the students could find the information they desired. However, they had difficulties with information overload. Also they are not sure whether the information resources are trustworthy and had difficulties to understand the information found.

In summing up, the findings indicate that:

- Information seeking consumes more time than expected by students.
- Computer Science postgraduate students encounter problems of information overload during the information seeking process.

- They do not have any problem in using appropriate keyword (Table 4.3). This shows that even though the satisfaction with the information found is high, they still have difficulties to understand it.

4.2 Questionnaire 2: A Survey on Students' Learning Style Awareness

This questionnaire used VARK LS test to identify students' preference (Fleming, 2010). The objectives of the questionnaire are to investigate students' awareness of LS and to identify students' LS preferences for evaluation purpose.

4.2.1 Results

Table 4.6 shows the distribution of the respondents' demographic profiles. The distribution of respondents is almost the same in gender and nationality. Most of the respondents (63.1%) were Masters by Coursework students.

Table 4.6: Distribution of the Respondent's Demographic Profile

Demographic Profile (N=111)		**Responses N (%)**
Gender	Male	57 (48.6%)
	Female	54 (51.4%)
Level of Study	Masters by Coursework	70 (63.1%)
	Masters by Research	9 (8.1%)
	PhD	32 (28.8%)
Nationality	Malaysian	52 (46.8%)
	Non-Malaysian	59 (53.2%)

a. LS Awareness

From Figure 4.6, we can see that half of the respondents (52.3%) were not aware of the term LS. The relationship between demographic profile and respondents' awareness of LS was examined using cross tabulation as shown in Table 4.7.

Female respondents are more familiar with the LS term with 54.4%. The most striking result to emerge from the table is that Malaysian students showed the highest result with 67.3%.

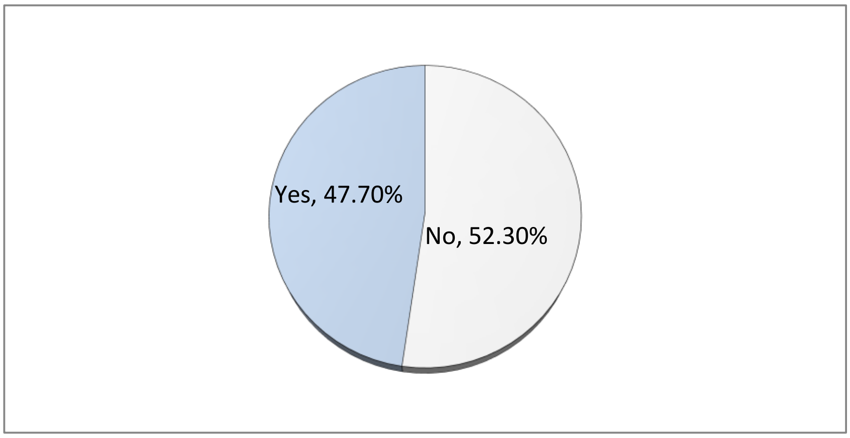


Figure 4.6: LS Awareness

Table 4.7: Students' Awareness of Their LS Based on Demographic

Demographic Profile (N=111)		**Awareness (%)**		
		Yes	**No**	**Total**
Gender	Male	40.7	59.3	100
	Female	54.4	45.6	100
Level of Study	Masters by Coursework	50	50	100
	Masters by Research	55.6	44.4	100
	PhD	40.6	59.4	100
Nationality	Malaysian	67.3	32.7	100
	Non-Malaysian	30.5	69.5	100

After the students were given an explanation about LS, they were asked if the awareness of LS is important in improving learning ability. Figure 4.7 presents the results obtained from the survey.

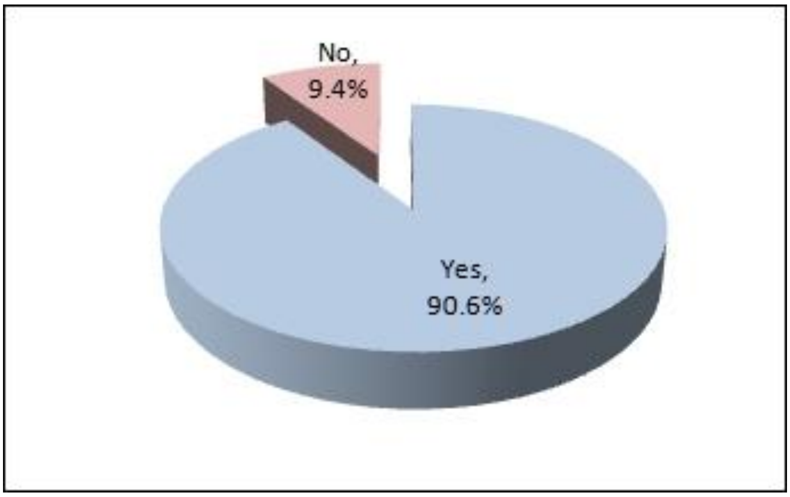


Figure 4.7: Awareness of the Importance of LS after Explanation

From Figure 4.7, it is apparent that most respondents believed that knowing LS is crucial in improving learning ability. Some respondents said that they know what LS is but they need more information about LS to understand (*S13*) and use it (*S85*).

b. Learning Style Preference

In this part, respondents were asked about their LS preference. Then, they had to answer VARK LS test.

1. LS preference based on respondents' perception

The chart in Figure 4.8 below shows the breakdown of respondents' LS preference. It is apparent from this figure that over half of the respondents (50.5%) perceived they have *Visual* preference. Only 7.2% believed they have *Aural* preference.

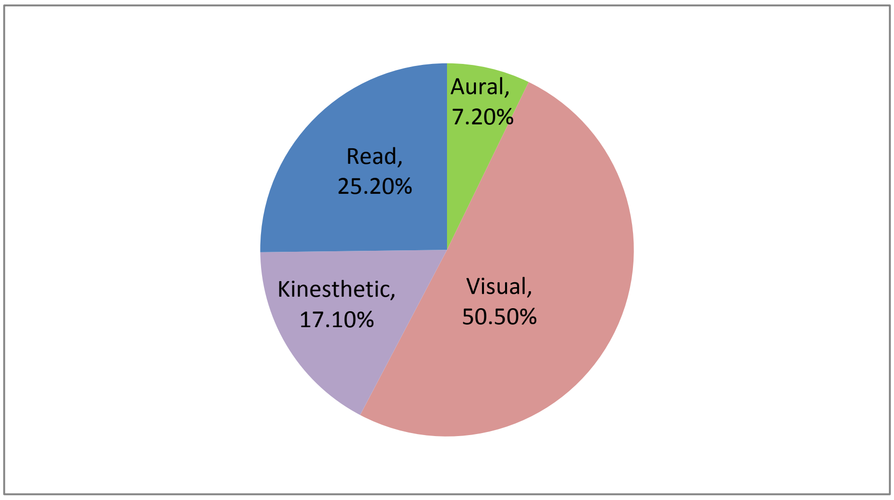


Figure 4.8: LS Preference Based on Respondents' Perception

2. LS preference based on LS test result

Figure 4.9 illustrate the LS preference based on VARK's LS test. The results show an even distribution of LS preferences which is different from Figure 4.8. *Multimodal* preference is added when the highest total of answer is the same for two or more preferences.

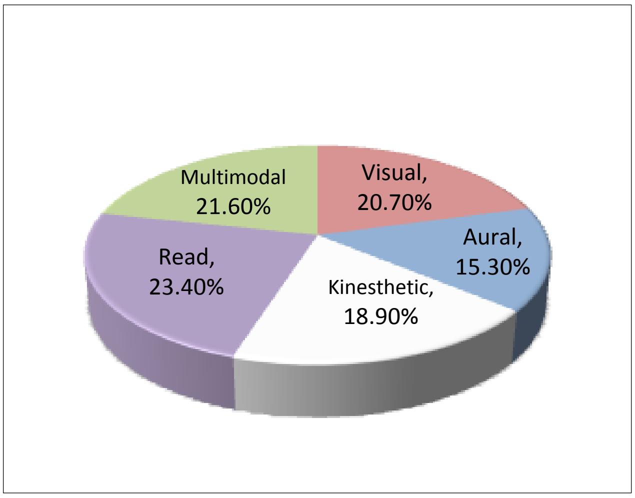


Figure 4.9: LS Preference Based on VARK LS Test Result

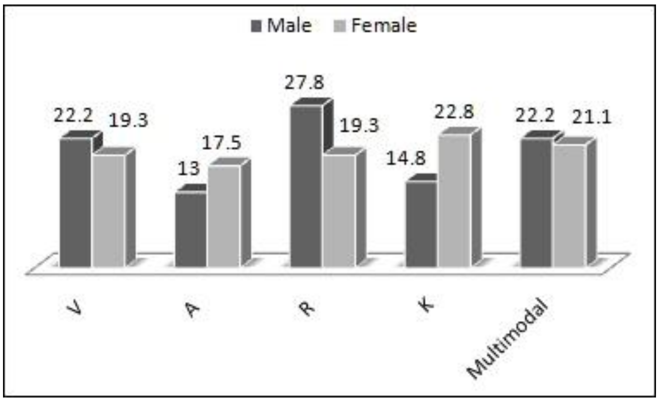


Figure 4.10: Male and Female LS Preference Based on VARK LS Test

Figure 4.10 illustrates LS preference based on gender. The highest preference for Male respondents is *Read* preference while Female respondents are *Multimodal* preference. Chi-square analysis shows that there are no significant difference between Male and Female ($\chi^2 = .23$, $p > .05$).

Results between respondents’ perception and actual test are compared. Only 37 out of 111 respondents' perceived and actual LS results were the same. 66.7% of the respondents produced different results. Figure 4.11 below shows the difference between perceptual and actual test of LS results.

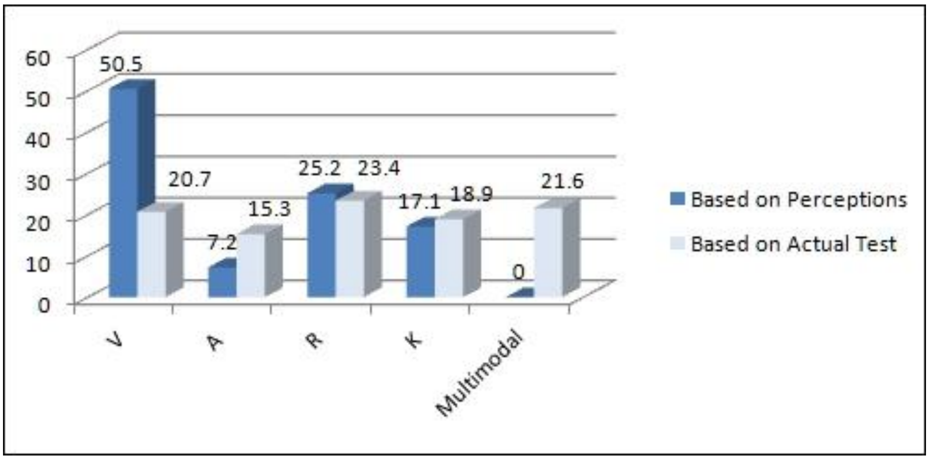


Figure 4.11: Comparison of LS preference: Perception versus VARK LS Test

From Figure 4.11, we can see that there are major difference in *Visual* (V), *Aural* (A) and *Multimodal.* The obvious difference between the perceived and actual LS results is shown in the V group. About half of the respondents (50.5%) initially believed that they are *Visual* but the VARK test revealed only 20.7% of them are. The test also revealed that 21.6% of the respondents are *Multimodal*. For detailed comparison see Table 4.8.

Table 4.8: Breakdown of Perceived Versus Actual LS Results

Perceived LS	**Results (%)**	**Actual LS**	**Results (%)**
Visual	50.5	Visual	14.4
		Audio	5.4
		Read/Write	9
		Kinesthetic	8.1
		Multimodal	13.6
Audio	7.2	Visual	1.44
		Audio	1.44
		Read/Write	1.44
		Kinesthetic	1.44
		Multimodal	1.44
Read/Write	25.2	Visual	2.7
		Audio	5.4
		Read/Write	10.8
		Kinesthetic	2.7
		Multimodal	3.6
Kinesthetic	17.1	Visual	1.8
		Audio	2.7
		Read/Write	3.6
		Kinesthetic	6.3
		Multimodal	2.7

Table 4.8 shows the breakdown of the perceived results based on the actual test results. For example, from the 50.5% of respondents who initially believed that they have *Visual* preference, only 14.4% actually belong to the *Visual* group. 13.6% of them in reality have *Multimodal* preference, 9% actually belong to the *Read/write* group, 8.1%

are of the *Kinesthetic* type and only 5.4% have *Aural* preferences. The results show that most of the respondents' perception is not matched with the actual result with 66.7%.

4.2.2 Discussion

In this section, we explore respondents' awareness toward LSs. We also compare the result between respondents' LS preference with their actual LS based on VARK LS test.

1. Respondents' awareness of LS

From the results, we can see about half of the respondents (52.3%) are unaware of their own LS preferences. This is in line with the findings from Rogers (2009) and Hassan et al. (2010) that shows some students are either not aware of LS or do not know their LS preference. Some of them know about LS, but do not take advantage of it in their learning process especially when it comes to choosing RMs. This leads to the mismatches between the LS of students and the presentation of the information in RM.

2. Respondents' LS preference

When asked to choose single sensory modality, most choose *Visual* preference (50.5%). Only 7.2% choose *Aural* preference. The perception of respondents on their LS is, however, not completely accurate. This is shown by the results from the actual LS test. Figure 4.9 illustrates the LS preference based on VARK's LS test. The results show an even distribution of LS preferences which is different from Figure 4.8. *Multimodal* preference is added when the highest total are the same for two or more preferences. A comparison of the two figures (Figure 4.8 and Figure 4.9) supports the findings that students are unaware of their own LSs that have been discussed above. The result is similar with Hassan et al. (2010), where the respondents' preferences do not match their' actual LS. The respondents in Hassan et al. (2010) also indicate that most of them think they are *Visual* preference with 53.13% while the actual test shows only 6.25%. The result also shows similarity for the *Aural* preference where it was the least preference selected by students with 6.26% while the actual test is 37.5%. This is also supported by results from Dobson (2010) where the results show the highest preference is *Visual* and the least was *Aural*.

From gender perspectives, the results show that there are no significant difference between Male and Female. This is in line with the result from Alkhasawneh et al. (2008), Dobson (2010), and Hassan (2009). The results are also similar to Geist and King (2008) and Reid (1987), i.e. Male prefers *Visual* and Female tends to be *Aural*. However, our findings show that Male prefers *Read* more than Female unlike James et al. (2011) and Dobson (2010). This could be due to our sample consisting of postgraduate students who tend to like reading.

4.3 Summary

This section discusses the results from both questionnaires and identifies whether the results can address the research questions. From Questionnaire 1, we can conclude that students face difficulties with:

- Information overload
- Understanding the information found
- Knowing where to find relevant information
- Categorizing their information needs

These results answer research question 1, which was ‘Do students have difficulties finding suitable RMs for their learning needs?’ It is hard for students to find suitable information when there is too much information available. Other issues are beyond the focus of this research.

Results from Questionnaire 2 show that most students were unaware of the term LS. They were also unaware of their own LS preference. These findings answer research question 3, which was ‘Are students aware of their LS?’ Not knowing their own LS might affect their ability to select suitable material for their learning needs.

Finally, we can summarize the findings from both questionnaires to the following two aspects:

- Difficulty in finding suitable RM due to information overload.
- Unawareness of the term LS and their LS preference.

CHAPTER 5

MAPPING READING MATERIAL ONTO STUDENTS' LEARNING STYLE

In this chapter we discuss how RM can be mapped onto students' LS. This leads to the development of our proposed tool namely *Learning Style based Information Seeking Tool* (LSIST). It begins with a summary of the problems students faced in finding the suitable RM, followed by the solution undertaken that explains how the tool is derived.

5.1 Mapping Process

The proposed IST, namely LSIST, retrieves RM based on students' LS preference. This tool maps RM onto LS. The mapping of the different types of RM onto the appropriate LS requires the RMs to be classified into suitable categories associated with the LS preferences. A suitable model for the LS is chosen for the purpose. Only then the classified RM can be mapped onto the LS preferences. Figure 5.1 shows the processes involved for the mapping of RM onto LS.

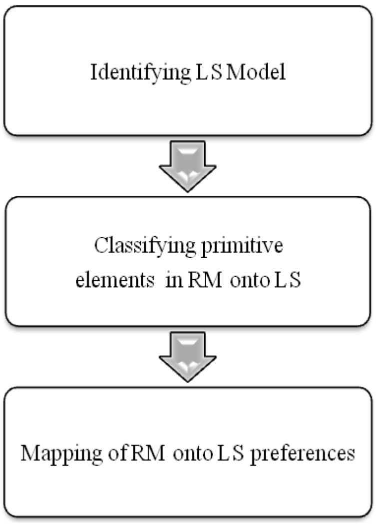


Figure 5.1: Mapping RM onto LS

5.2 Identifying Learning Style Model

VARK LS model by Fleming and Mills (1992) is chosen because the distinguishing elements that delineate each of the four LS preference as described in Table 5.1 can be mapped to the primitive elements in RM. The primitive elements are listed in Table 5.2. Hence this is the most suitable model to be used for the classification of primitive elements in RM. LS preference is identified using VARK LS test which consists of sixteen multiple choice questions with four answer selections corresponding to the four preferences as presented in Appendix B (Fleming, 2010).

Table 5.1: VARK LS Preferences

LS Preference	Description
Visual	Visual preference prefers the use of diagrams, pictures, slides, graphs and flow charts to represent printed information.
Aural	This perceptual mode describes a preference for information that is heard or spoken.
Read/write	Read/write preference prefers printed words and text as a means of taking in information.
Kinesthetic	Kinesthetic preference refers to learning achieved through experience and practice.

In this research, we only consider three types of preferences which are Visual, Read/write and Kinesthetic. Aural preference is ignored because the RMs considered were not in audio form. However, those with Aural preference can still use this proposed tool because it can still classify their alternative preferences.

5.3 Primitive Elements in Reading Material

To map RM onto students' LS, primitive elements in the RM need to be identified and classified into suitable categories corresponding to the respective LS preference.

5.3.1 Identifying Primitive Elements in Reading Material

Table 4.2 below lists the various primitive elements in RM that need to be identified for the purpose of categorization. The primitive elements are identified from LOM and other studies done by Antoine et al. (1992), Bayer et al. (1992), Dori et al. (1997), Aumann et al. (2006) and Lu et al. (2006).

Table 5.2: List of Primitive Elements

Primitive Elements	Description
Text	Text consists of letters alphabet, character and symbols and can be presented in word, line, column and paragraph (Bayer et al., 1992).
Characters	Characters are usually small connected components lined up to make words (Antoine et al., 1992).
Symbol	A mark or character used as a conventional representation of an object, function, or process (Oxford, 2005).
Word	Word is a frame containing a group of one or more aligned characters separated by white space (Dori et al., 1997).
Line	A line of text (Dori et al., 1997).
Column	A column of text (Dori et al., 1997).
Paragraph	A delimited block of text comprising a paragraph (Dori et al., 1997).
Picture	A visual representation or photographed (Oxford, 2005).
Screenshot	An image of the display on a computer screen (Oxford, 2005).
Map	A diagrammatic representation of an area of land or sea showing physical features, cities, and roads (Oxford, 2005).
Histogram	A diagram consisting of rectangles whose area is proportional to the frequency of a variable and whose width is equal to the class interval (Oxford, 2005).
Bar Chart	A diagram in which the numerical values of variables are represented by the height or length of lines or rectangles of equal width (Oxford, 2005).
Pie Chart	A type of graph in which a circle is divided into sectors that each represent a proportion of the whole (Oxford, 2005).
Table	A set of facts or figures systematically displayed, especially in columns (Oxford, 2005).
Equations	An equation, in a mathematical context, is generally understood to mean a mathematical statement that asserts the equality of two expressions.
Notation	A series or system of written symbols used to represent numbers, amounts, or elements in mathematics (Oxford, 2005).
Formula	A mathematical relationship or rule expressed in symbols (Oxford, 2005).
Flowchart	A diagram of the sequence of movements or actions of people or things involved in a complex system or activity (Oxford, 2005).
Diagram	Diagrams are images that show arrangements and relational dependencies among a series of components (Lu et al., 2009).
Network diagram	A schematic depicting the nodes and connections amongst nodes in a computer network or, more generally, any telecommunications network.

5.3.2 Categorization of Primitive Elements

From the list in Table 4.2, the primitive elements in RM are divided into two categories: *Text* and *Non-text*.

- *Text* category consists of letters alphabet, character and symbols and can be presented in word, line, column and paragraph (Bayer et al., 1992).
- *Non-text* category comprises the rest including images that are labeled as figures within the embedding document (Dori et al., 1997).

Then, all the primitive elements under the *Non-text* category is further classified into four sub-categories which are photograph, graphic, semi graphic and diagram. The sub-categories are derived from Dori et al. (1997) and Lu et al. (2009). This is described in Table 5.3 and presented in Figure 5.2.

Table 5.3: *Non-Text* Category

Non-text Category	**Description**	**Examples**
Photograph	A photograph is a continuous-tone image recorded by a camera or created by photo processing software (Li & Gray, 2000).	Picture, screen shot and map
Graphic	Graphics can be 2D or 3D plots. They are defined in a 2D or 3D coordinate system as series of points, lines, curves or areas that represent the variation of a variable with respect to another variable (Lu et al., 2009).	Scatter plot, histogram, bar chart and pie chart
Semi graphic	Semi graphics contain predominantly symbolic information whose meanings are determined by their spatial relationship to one another (Nagy et al., 1997)	Table, equation, notation and formula.
Diagram	Diagrams are images that show arrangements and relational dependencies among a series of components, such as rectangles, ovals, diamond, etc (Lu et al., 2009).	Flowchart, network diagram, Venn diagram and UML diagram.

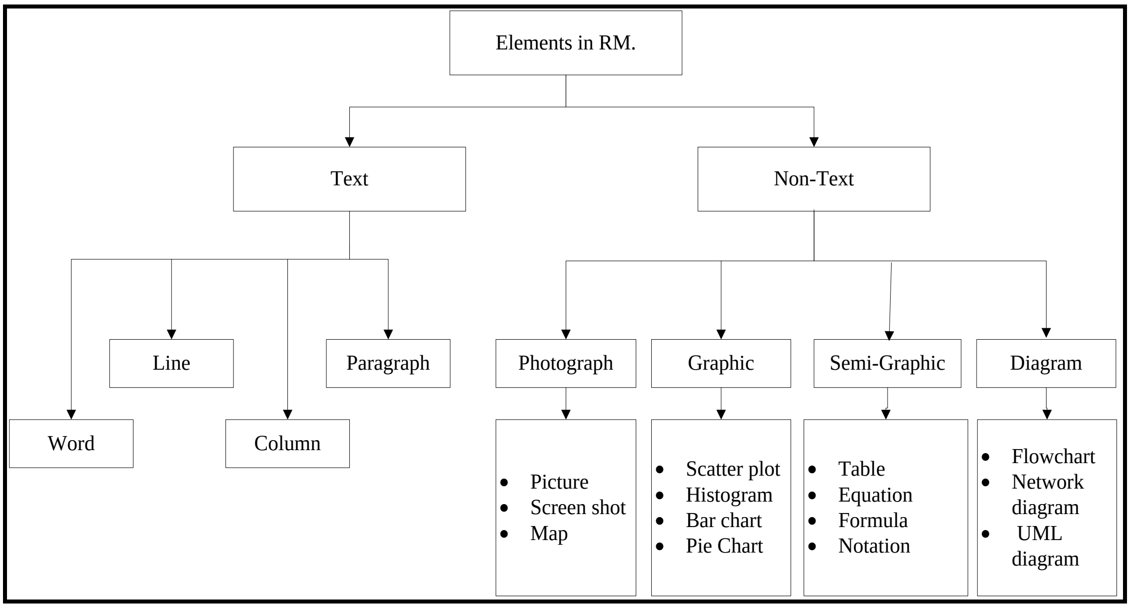


Figure 5.2: Primitive Elements in RM

5.3.3 Mapping Primitive Elements in Reading Material onto Learning Style

The two main categories described above namely the *Non-text* category and *Text* category map suitably to the Visual and Read/write preference respectively. The text category however can be further divided into Text (K) which consists of words such as examples, case studies, practice and applications. This category can then be mapped to the Kinesthetic preference. Text (R) category contains those words not belonging to the Text (K) category. The mapping of the primitive elements categories to the LS preference is shown in Table 5.4.

Table 5.4: Categorization of Primitive Elements According to LS Preferences

Primitive Elements		LS Preference
Category	**Sub-Category**	
Non-Text	Photograph, graphic, semi-graphic, diagram.	Visual
Text (K)	Words like examples, case studies, practice and applications.	Kinesthetic
Text (R)	All words except words describing Kinesthetic preference.	Read/write

Then, all primitive elements were standardized using identifier from Learning Resource Type in IEEE LOM as shown in Table 5.5. These identifiers will be used to identify the LS value in RM.

Table 5.5: Identifier based on LOM

LS Preference	Identifier
Visual	Figure, diagram, map, chart, graph, flowchart, arrow, circle, hierarchy, hierarchies, picture, table, equation, notation, formula, histogram, scatter plot, screenshot
Read/write	All words except words describing Visual and Kinesthetic preference
Kinesthetic	Example, practice, case study, exercise, simulation, experiment, self-assessment, application

As a conclusion, after the classification, all the primitive elements in RM can now be mapped onto the LS preferences as shown in

Figure 5.3.

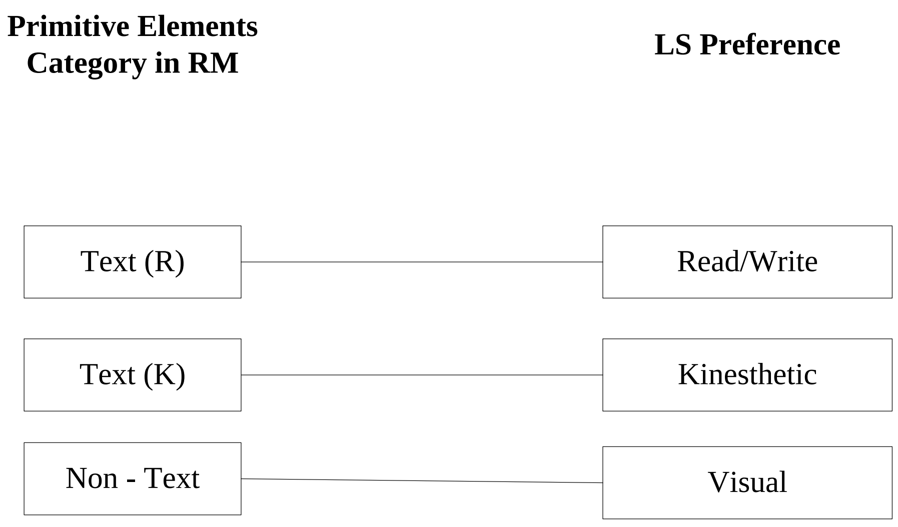


Figure 5.3: Mapping Primitive Elements to LS Preference

5.4 Mapping of Reading Material onto Learning Style Preferences

In order to map RM onto LS preferences, a classification model is constructed. The model consists of five processes as shown in

Figure 5.4.

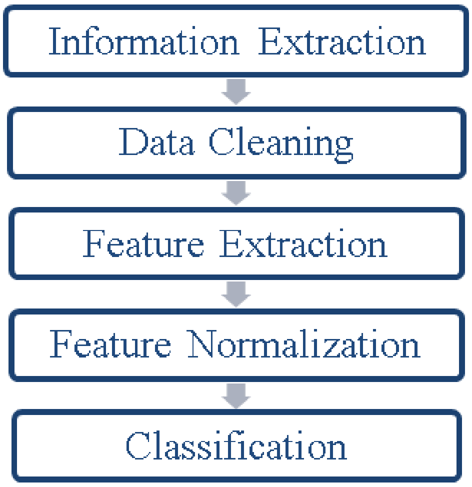


Figure 5.4: Mapping RM onto LS Preferences Processes

5.4.1 Information Extraction

The RMs used in this study are limited to PDF type only. The content in RM is extracted using iText; a PDF extraction tool as discussed in section 2.2.2. Output from the information extraction is a page with text content and embedded images.

5.4.2 Data Cleaning

After the content of a PDF file is extracted, data cleaning process is carried out. Data cleaning prepares the content for the feature extraction process by removing outliers as it can be incorrectly identified as features. Hence, unrelated content which is not identified as words is removed. This prevents unreliable output due to incorrectly identified object. Data discretization is then carried out to replace raw data values by ranges or higher conceptual levels (Han & Kamber, 2000).

The data cleaning process is listed below and a sample of its pseudo code is presented in Figure 5.5.

- **Removing Outliers**

  Outliers are non-words characters, i.e. any character apart of a-z, A-Z, 0-9. All outliers are removed from the content.

- **Data Discretization**

  Data discretization transforms raw data into specific labels. For example, a word is counted as a word if it contains more than two (>2) characters.

```
if (string.IsNullOrWhiteSpace(removed)) break;
                    removed = removed.Replace(s.ToString(), string.Empty);
                }
                string[] words = content.Split(' ');
                if (words.Length > 2 && !string.IsNullOrWhiteSpace(removed))
                    readWritePages++;
                }
```

Figure 5.5: Pseudo Code for Data Cleaning

5.4.3 Feature Extraction

This is a process to transform the input data (RM) into a reduced representation set of features called feature vector. In this work, there will be three vectors representing the features associated with the three LS preferences namely Visual, Read/Write and Kinesthetic. To produce the three feature vectors of a RM, each feature vector needs to be identified and calculated.

- **Identifying Feature Vector**

RMs are represented as collections of pages containing text and images. Feature extraction is conducted at the page level by identifying the respective identifiers of each feature vector as listed Table 5.6. Feature vector value for RM is the sum of identified feature vectors in the whole RM.

The existence of the *Read/Write* and *Kinesthetic* feature vectors can be identified by extracting words referring to the identifiers in Table 5.6. For *Visual* feature vector, iText can only extract images embedded inside PDF. If the primitive elements are objects in the form of vector images, they are not considered as images as they cannot be detected. Vector images are combinations of arrow, line and word which represent the table, flow, chart and other object. To extract vector image as image in PDF file, knowledge in image analysis is needed. However, object identification and image analysis is beyond the focus of this study.

This study used logical text associated to the identifiers of the *Visual* feature vector to extract the *Non-text* category of primitive elements. Identifiers for the *Visual* feature vector are divided into three groups:

- **Text (Va)**

  Identifiers that represent Photograph, Graphic and Diagram categories

- **Text (Vb)**

  Identifiers that represent Semi graphic category

- **Non-Text**

  Object detected as image embedded in RM content

Each feature vector is identified using identifier tags that represent all identifiers describing LS preferences as shown in Table 5.6.

Table 5.6: Feature Vectors and Its Identifier

Feature Vector	Identifier Tag	Identifier
Visual	Text (Va)	Identifier describing Visual preference: *Figure, diagram, map, chart, graph, flowchart, arrow, circle, hierarchy, hierarchies, picture, histogram, scatter plot, screenshot*
	Text (Vb)	Identifier describing Visual preference: *Table, equation, notation, formula*
	Non-Text	Image embedded in PDF file
Read/Write	Text (R)	All words except words in the *Visual* and *Kinesthetic* feature vector with number of words greater than 10
Kinesthetic	Text (K)	Identifier describing Kinesthetic preference: *Example, practice, case study, exercise, simulation, experiment, self-assessment, application*

- **Feature Vector Calculation**

Feature vector is calculated for each page. Identifiers in Table 5.5 are used in identifying the LS value in RM. All the identifiers are given weights as shown in Table 5.7. Identifiers describing the Semi graphics category are given weights of 0.5 because they are not completely considered as visual and need more effort to understand them (Nagy et al., 1997).

Table 5.7: Identifiers Value

Feature Vector	Identifier	Weighting
Visual	Figure	+1
	Diagram	+1
	Map	+1
	Chart	+1
	Graph	+1
	Flowchart	+1
	Arrow	+1
	Circle	+1
	Hierarchy	+1
	Hierarchies	+1
	Picture	+1
	Histogram	+1
	Scatter plot	+1
	Screenshot	+1
	Table	+0.5
	Equation	+0.5
	Notation	+0.5
	Formula	+0.5
Read/write	All words except words describing Visual and Kinesthetic preference	+1
Kinesthetic	Example	+1
	Practice	+1
	Case study	+1
	Exercise	+1
	Simulation	+1
	Experiment	+1
	Self-assessment	+1
	Application	+1

For each feature vector, the maximum value for each page is one. If an identifier for any feature vector is found on the page then the value of that feature vector is increased accordingly to the identifier tags representing the identifiers. If the value of that feature vector is already one then the next identifier of the same feature vector found on the same page is ignored.

The calculation of the *Visual*, *Read/write*, and *Kinesthetic* feature vectors value for a page is given below:

To calculate the *Visual* feature vector *(v)* value of a page,

$$\begin{cases} v = 1, & if\ Text\ (Va) + Text(Vb) + Non - Text \geq 1 \\ v = 0.5, & if\ Text\ (Vb) \geq 1 \\ v = 0, & otherwise \end{cases}$$

To calculate the *Read/write* feature vector (*r)* value of a page,

$$\begin{cases} r = 1, & Text\ (r) \geq 1 \\ r = 0, & otherwise \end{cases}$$

To calculate the *Kinesthetic* feature vector (*k)* value of a page,

$$\begin{cases} k = 1, & Text\ (k) \geq 1 \\ k = 0, & otherwise \end{cases}$$

Note that, values for the whole document are used instead of the number of identifiers in the page. The results are shown in percentage instead of categorizing each RM onto a specific LS preference. This is to ensure any combination of LS preference eg. Aural and Multimodal can be catered for.

The total of all three feature values for the RM is calculated as such:

$$t = \sum v + \sum r + \sum k$$

Where;

t = Total of all feature vector values

v =Value of the *Visual* feature vector on each page

r = Value of the *Read/write* feature vector on each page

k = Value of the *Kinesthetic* feature vector on each page

The calculation for the three feature vectors for the RM is given below:

Value for *Visual* feature vector, *v*,

$$v = \left(\frac{\sum v}{t}\right) \times 100$$

Value for *Read/write* feature vector, *r*,

$$r = \left(\frac{\sum r}{t}\right) \times 100$$

Value for *Kinesthetic* feature vector, *k*,

$$k = \left(\frac{\sum k}{t}\right) \times 100$$

Below is an example of the calculation for the feature vectors for a PDF file with two pages.

Page 1: contains two Visual identifiers, a few text (Read/write) and zero Kinesthetic identifier.

For this page, the *Visual*, *Read/write* and *Kinesthetic* feature vector values will be 1, 1 and 0 respectively:

$$v \leftarrow 0, v \leftarrow v + 1 = 0 + 1 = 1$$

$$r \leftarrow 0, r \leftarrow r + 1 = 0 + 1 = 1$$

$$k \leftarrow 0, k \leftarrow k + 0 = 0 + 0 = 0$$

Page 2: contains zero Visual identifier, a few text (Read/write) identifiers and zero Kinesthetic identifier.

The *Visual*, *Read/write* and *Kinesthetic* feature vector values of page 1 are added to the Visual, Read/write and Kinesthetic feature vector values of page 2 resulting in 1, 2 and 0 respectively:

$$v \leftarrow 1, v \leftarrow v + 0 = 1 + 0 = 1$$

$$r \leftarrow 1, r \leftarrow r + 1 = 1 + 1 = 2$$

$$k \leftarrow 0, k \leftarrow k + 0 = 0 + 0 = 0$$

The total of all three feature vector values for the whole document *(t)* is:

$$t = v + r + k = 1 + 2 + 0 = 3$$

Then each feature vector, *v, r* and *k* is calculated respectively:

$$v = \left(\frac{1}{3}\right) \times 100 = 33.33\%$$

$$r = \left(\frac{2}{3}\right) * 100 = 66.77\%$$

$$k = \left(\frac{0}{3}\right) * 100 = 0\%$$

LS value for this RM is 33.33% for *Visual*, 66.77% for *Read/write* and 0% for *Kinesthetic*. The results are shown as three different percentages representing the three LS preferences instead of categorizing each RM onto a specific LS preference. This is to ensure any combination of LS preference eg. Aural and Multimodal can be catered for. A sample pseudo code to accumulate the *Read/write* feature vector value is presented in Figure 5.6.

```
r ← 0                                  . reset readwrite counter
for i = 1 → doc.pages do
    w ← doc.page(i)                    . get doc content at page i
    if w ≥ 2 then                      . when doc has more than 2 words
        r ← 1 + r                      . increment readwrite counter
    end if
```

Figure 5.6: Sample Pseudo Code for Feature Extraction Algorithm

5.4.4 Feature Vector Normalization

The feature vector values of the students and RM are not standardized. Standardization removes scale effects caused by the use of features with different measurement scales. The features involve in this study are:

- **RM feature ($\boldsymbol{R}$)**

  Values for R are taken from values of *v, r and k* discussed in section 5.4.3.

- **Student feature ($\boldsymbol{S}$)**

  Values for S are derived from the VARK LS test result representing Visual (v), Read/write (r) and Kinesthetic (k) preference. VARK LS test consists of sixteen questions so the result will be in the range one to sixteen. For example, a student with test result v=10, r=9 and k=14 indicates that the student is a Kinesthetic learner.

Values for R are in percentage $[0,100]$ whilst values for S are in the range of $[1,16]$. To standardize these two values they are normalized to $[0,1]$. Once standardization is performed on a set of features, the range and scale should be similar, provided the distributions of raw feature values are alike. For the normalization, we use Softmax formula. Softmax formula is chosen because it can ensure all of the output values are between 0 and 1 and that their sum is 1 (Bridle, 1990).

To normalized the feature vector for RM (R),

$$R_c = \frac{Q_c}{\sum_{j=1}^{n} Q_j}$$

Where,

R_c = Value of RM feature vector c after normalization [0,1]

Q_c = Value of feature vector c

$c \in \{visual, read, kinesthetic\}$

Q_j = Value of feature vector *of j*

n = Number of feature vector

j = 1,2,..,n

Below is an example of the calculation for the RM feature vectors. The initial values for RM feature vectors are taken from previous example of feature extraction calculation whereby their initial values are:

$$v \leftarrow 33.33$$

$$r \leftarrow 66.67$$

$$k \leftarrow 0$$

By using the above formula, we calculate the value for RM feature vectors:

$$R_v = \frac{Q_v}{Q_v+Q_r+Q_k} = \frac{33.33}{33.33+66.67+0} = \frac{33.33}{100} = 0.33$$

$$R_r = \frac{Q_r}{Q_v+Q_r+Q_k} = \frac{66.67}{33.33+66.67+0} = \frac{66.67}{100} = 0.67$$

$$R_k = \frac{Q_k}{Q_v+Q_r+Q_k} = \frac{0}{33.33+66.67+0} = \frac{0}{100} = 0$$

After normalization, total for RM feature vectors = $R_v + R_r + R_k$ = 0.33 + 0.67 + 0 = 1

Value for Student feature vectors are between zero and sixteen. To normalize feature vector for Student (S),

$$S_c = \frac{Q_c}{\sum_{j=1}^{n} Q_j}$$

Where,

S_c = Value of student feature vector *c* after normalization [0,1]

Q_c = Value of feature vector *c*

$c \in \{v, r, k\}$

Q_j = Value of feature vector *of j*

n = Number of feature vector

j = 1,2,..,n

Below is an example of the calculation for the Student feature vectors. The initial value for Students feature vectors are:

$$v \leftarrow 10$$

$$r \leftarrow 9$$

$$k \leftarrow 14$$

By using the above formula, we calculate the value for Student feature vectors:

$$S_v = \frac{Q_v}{Q_v + Q_r + Q_k} = \frac{10}{10 + 9 + 14} = \frac{10}{33} = 0.303$$

$$S_r = \frac{Q_r}{Q_v + Q_r + Q_k} = \frac{9}{10 + 9 + 14} = \frac{9}{33} = 0.273$$

$$S_k = \frac{Q_k}{Q_v + Q_r + Q_k} = \frac{14}{10 + 9 + 14} = \frac{14}{33} = 0.424$$

After normalization, total for Student feature vectors

$= S_v + S_r + S_k = 0.303 + 0.273 + 0.424 = 1.$

5.4.5 Classification Using k-Nearest Neighbor

To map RM with LS preference, k-Nearest Neighbor (k-NN) method is used in the Matching algorithm. K-NN is a method for classifying objects based on closest relations in the feature space. This method is used because we want to consider all feature vector values without categorizing each document to only a specific LS preference. This reduces information loss.

The k-NN classifier is based on the Euclidean distance. Euclidean distance is used to calculate distance between students' LS (*S*) and RMs' LS (*R*) value. It is chosen because it is sufficiently accurate to calculate the distance between multiple feature vector to get the relationship (Danielsson, 1980). Distance between feature vectors from one to another is calculated to check the closest relationship. In this case, how close the LS feature vector of *S* and *R* is related.

Feature vectors for both features are *v*, *r* and *k* values which have been normalized to [0, 1] in section 5.4.4 are used. We plot the feature vector of *S* and *R* as coordinate in Cartesian coordinate. Relation *d(S,R)* between *S* and *R* LS value is calculated from the distance value of LS feature vector. The lower value means the closest relation.

$$d(S,R) = \sqrt{(S_v - R_v)^2 + (S_r - R_r)^2 + (S_k - R_k)^2}$$

Where;

d = Distance

S = Student

R = RM

v = *Visual* feature vector

r = *Read/write* feature vector

k = *Kinesthetic* feature vector

From the results, all the RMs value is then ranked in ascending order.

5.4.6 Matching Reading Material onto Students' Learning Style

The process of matching RM onto students' LS is executed when student enter keyword for the query. The process for searching the suitable RM is described as follows and the Matching Algorithm is presented in Figure 5.7

Step1: Query keyword of RM

Step 2: For each RM with similar keyword with query,

1. Calculate LS value distance between RM and student (See 5.4.5)
2. Save distance and RM to database
3. Sort the result from lowest to highest distance

Step 3: Show RM from the database

```
procedure SEARCHDOC(query)  ⊳ sort by student and RM matching
    rank ← hashtable
    Sv ← student.visual
    Sr ← student.readwrite
    Sk ← student.kineshtetic
    for i = 1 → docs.length do
        if RM(i).title ≡ query then
            Rv ← RM(i).visual
            Rr ← RM(i).readwrite
            Rk ← RM(i).kineshtetic
                                                  - calculate VARK distance
            distance = √((Sv − Rv)² + (Sr − Rr)² + (Sk − Rk)²)
            rank ← distance, RM(i)
        end if
    end for
    sort(rank)                                    ⊳ sort by the lowest distance
    return rank
end procedure
```

Figure 5.7: Pseudo Code for Matching Algorithm

5.5 Summary

This chapter discussed the three main processes involved in mapping RM to students' LS namely identification of the LS model, classification of primitive elements and mapping RM to LS preference. To map RM to LS preference, the process of information extraction and data cleaning, feature extraction, feature normalization and RM classification were also explained. The next chapter presents the design and development of LSIST comprising those processes as its components.

CHAPTER 6

DEVELOPMENT OF LSIST PROTOTYPE

A *Learning Style Based Information Seeking Tool* (LSIST) is developed for this research. The development process involves the system design and its implementation. The architecture of LSIST is discussed followed by system design and implementation.

6.1 LSIST Architecture

LSIST architecture involves two types of users; administrators and students. Administrators provide RM to the tool. RM used as a source for this prototype is restricted to those formatted in PDF. LSIST architecture is discussed in section 3.3. For evaluation purpose, another module for *Keyword based Search* is added to the prototype to complement the *LS based Search* module as shown in

Figure 6.1.

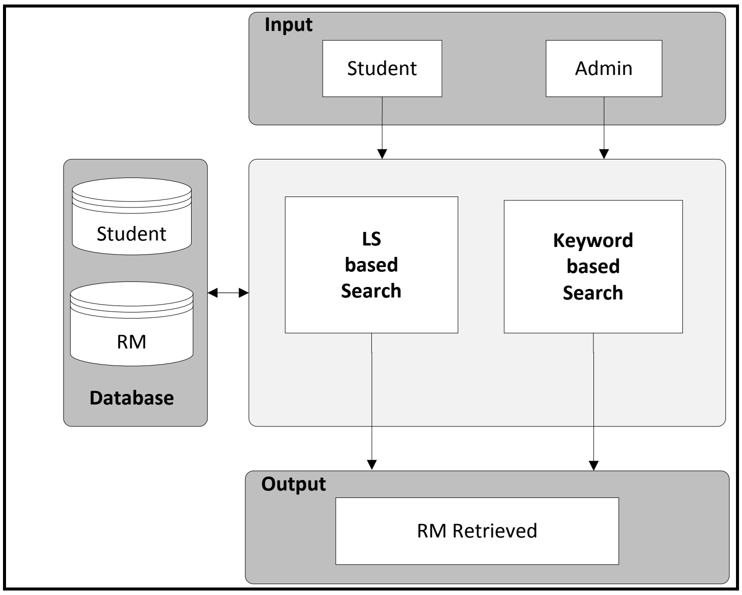


Figure 6.1: LSIST Architecture with *Keyword based Search* module

LSIST architecture consists of four main modules; *Input*, *LS based Search*, *Database* and *Output*:

1. *Input* Module

Input module receives inputs from *Student* and *Admin*. Students provide inputs such as students's profile, LS preferences and search query. These inputs are used to develop a *User Model* in the *LS based Search* module and stored in *Student Database*. *Input* module also receives RM profile such as title, author, and LS value from the administrator.

2. *LS based Search* Module

LS based Search retrieves RM based on keyword and LS. This module contains three main components; *User Model*, *RM Classification* and *Matching*.

- *User Model* - *User Model* receives information from *Input* module and stores it in *Database* module. In *User Model*, students are classified based on LS preference. The filtering is based on LS test. LS test result are then stored in *Student Database*.
- *RM Classification* - This component identifies LS value from RM using Feature Extraction algorithm as discussed in section 5.4.3 Feature Extraction and stores the information in *RM Database*.
- *Matching* - *Matching* component uses information stored in *Database* module to execute searching and matching process using Matching algorithm as discussed in section 5.4.5).

3. *Keyword Based Search* Module

This module is developed based on keyword search technique.

4. *Database* Module

This module comprises two databases which are *Student Database* and *RM Database*:

- *Student Database* stores student's information such as gender, level of study and LS preference.
- *RM Database* stores RM's information such as title, author and LS value.

5. *Output* Module

The recommended RM is displayed in *Output* module. Students are asked to give feedback regarding the RM retrieved from LSIST. They are required to rate the RM on a scale of 1 to 5.

The design for this prototype is discussed in the next section.

6.2 System Design Using SSADM

This system is developed using Structured Systems Analysis and Design Method (SSADM) (Weaver et al., 1998). It contains five modules which are feasibility study, requirements analysis, requirements specification, logical system specification and physical design.

6.2.1 Feasibility Study

In this module, information for scope and problem is identified from Chapter 2 and Chapter 4.

6.2.2 Requirement Analysis

Requirement Analysis module enables a full understanding of the requirement of the proposed system. The objective of this module is to define the boundaries and objectives of the proposed system. Deliverables for this module are:

- Current services description which includes current Entity Relationship Diagram (ERD) and Data Flow Diagram (DFD).
- Requirement Catalogue for the new system
- User Catalogue for the new system

First, current services description is studied to investigate the functionality, data and system users. ERD is identified to gain an initial understanding of the system and its interrelationships, while DFD is identified to understand the flows of data around the system. Existing ERD shows that the existing system only involves *Student* and *RM* entities. *Student* can search for many *RMs* and a *RM* can be retrieved by many *Students*. DFD (Figure 6.2) for existing system are shown below. In this study, existing system refers to the basic process of IST discussed in section 2.6.

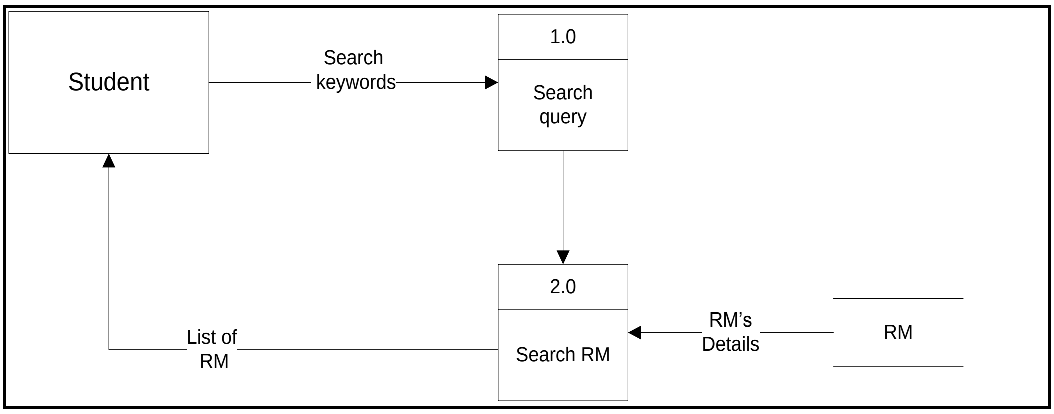


Figure 6.2: Existing DFD

Existing DFD shows that in the existing system, students insert keywords, then the system lists out all RM that contains the keywords. After the analysis of the existing system, Requirement Catalogue and User Catalogue for the proposed system are identified. Requirement Catalogue contains details which help to identify the requirement for reference purpose. Below is the Requirement Catalogue for the proposed system.

<table>
<tr><th colspan="3">Requirements Catalogue Entry</th></tr>
<tr><td colspan="3">Project Learning Style Based Information Seeking Tool Version 1</td></tr>
<tr><td colspan="3">Functional Requirement
The new system should provide search facilities that can help student find RM based on their LS</td></tr>
<tr><td colspan="3">Non-functional requirements</td></tr>
<tr><td>Description
Access restriction – only to registered students</td><td>Target value
Postgraduate students</td><td>Acceptable range
Student enrolling with Master and PhD</td></tr>
<tr><td colspan="3">Benefits
At present students can retrieve all RMs based on keyword. This means they can retrieve a lot of RMs. Using LS to narrow down the selection will help them find suitable RMs for their learning needs.</td></tr>
<tr><td colspan="3">Comments/suggested solutions
VARK LS model is used
Audio Material is not included</td></tr>
<tr><td colspan="3">Related Documents
Chapter 2, Chapter 3, and Chapter 4</td></tr>
</table>

Figure 6.3: Requirement Catalogue Entry

Then, a user catalogue (Table 6.1) is produced. The user catalogue describes system users and the tasks they perform.

Table 6.1: User Catalogue

Job Title	Activities Description
Student	Insert student details
	Answering VARK LS test
Registered Student	Insert keywords and search for RM
	Give feedback
Admin	Insert RM
	Perform extraction analysis on RM

6.2.3 Requirement Specification

This module specifies the requirement of the system to provide a clear guidance to design the system. Deliverables for this module are:

- Required ERD
- Required DFD
- User Roles
- User Roles/Function Matrix

Required ERD is shown in Figure 6.4.

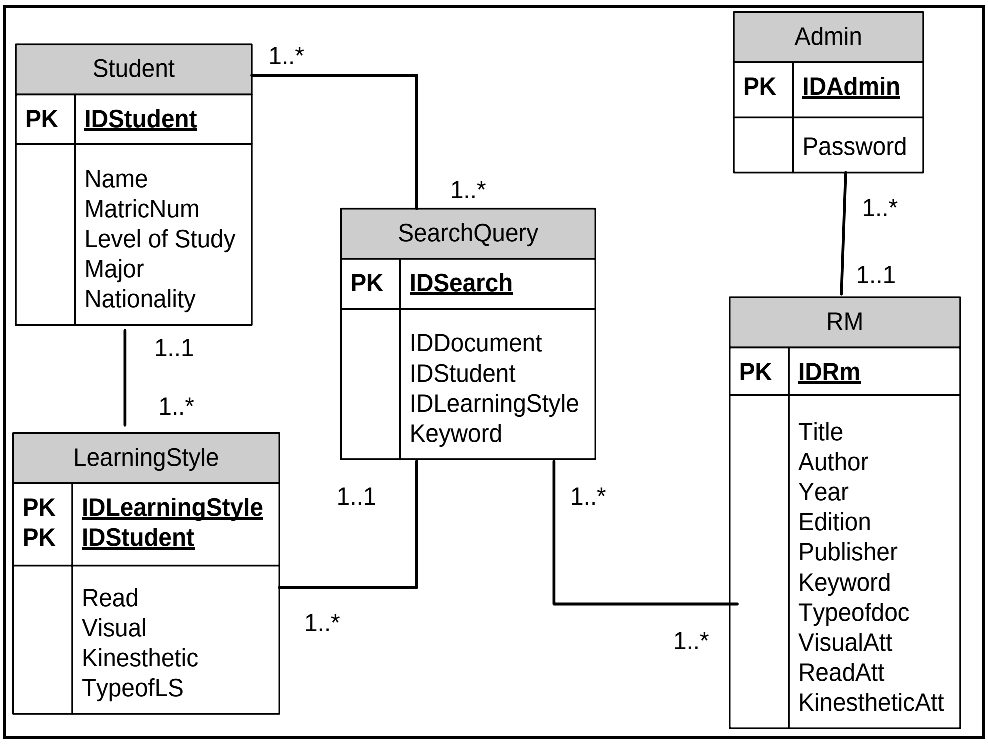


Figure 6.4: Required ERD

From the ERD above, a Required DFD is produced as shown in Figure 6.5.

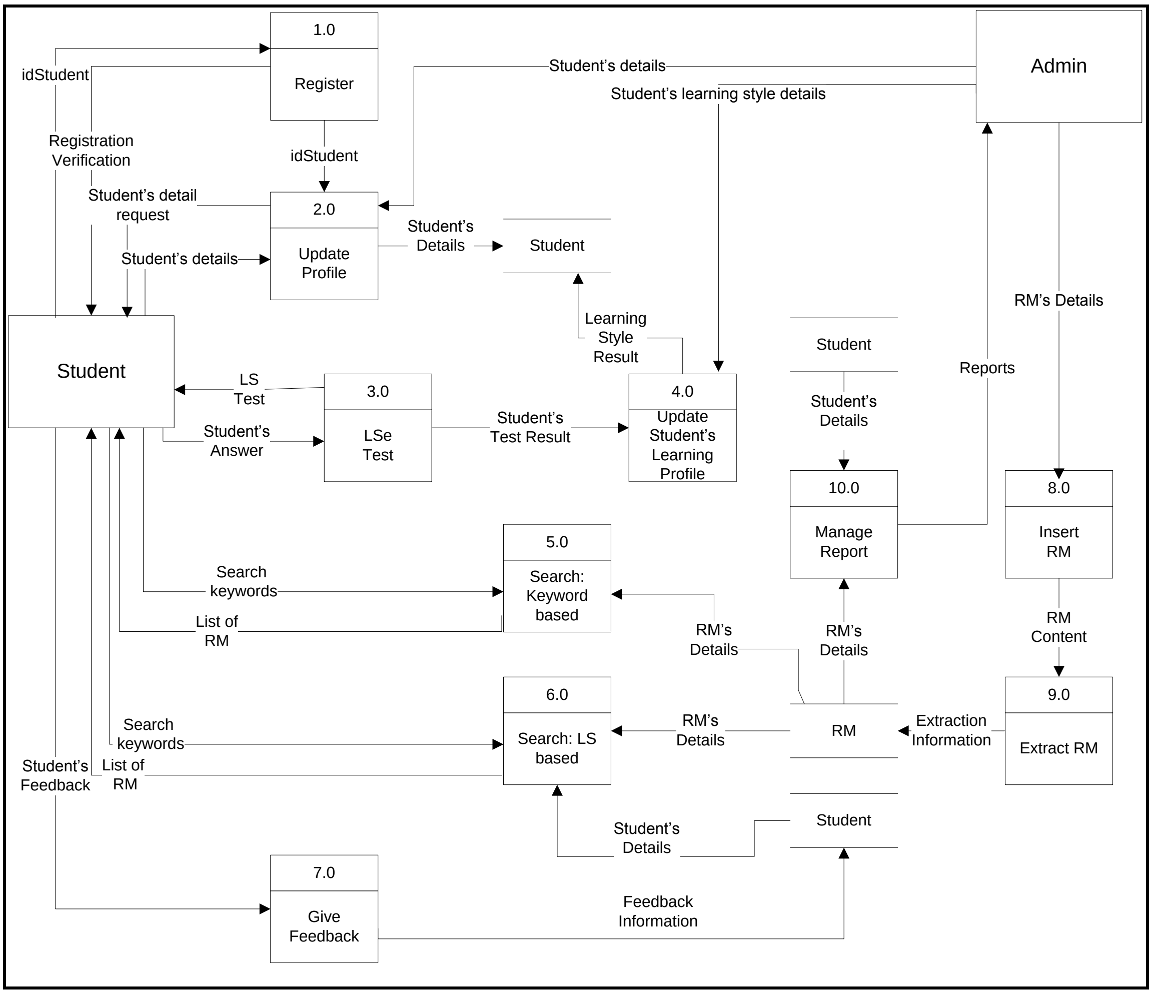


Figure 6.5: Required DFD

From the required DFD, we can identify external entity description (Table 6.2), Input/Output description (Table 6.3) and function and events (Table. 6.4).

Table 6.2: External Entity Description

Name	Description
Student	Search RM
Admin	Manage students and RM

Table 6.3: Input/Output Description for LSIST

From	To	Data Flow Name	Data Content
Student	1.0: Register	idStudent	Student identification number
1.0: Register	Student	Registration Verification	Confirmation of registration
Student	2.0: Update Profile	Student’s Detail Request	Request for student’s detail
2.0: Update Profile	Student	Student’s Detail	Student’s information
2.0: Update Profile	Student Database	Student’s Detail	Student’s information
Student	3.0: LS Test	LS Test	LS test request
3.0: LS Test	Student	Student’s Answer	Student’s LS test information
3.0: LS Test	4.0: Update Student’s Learning Profile	Student’s Test Result	Student’s LS test result information
4.0: Update Student’s Learning Profile	Student Database	LS result	Student’s LS test result information
Student	5.0: Search: Keyword based	Search keywords	List of keywords
RM Database	5.0: Search: Keyword based	RM’s Detail	RM’s information
5.0: Search: Keyword based	Student	List of RMs	List of RMs based on keyword
Student	6.0: Search: LS based	Search keywords	List of keywords
Student Database	6.0: Search: LS based	Student’s Detail	Student’s information
RM Database	6.0: Search: LS based	RM’s Detail	RM’s information
6.0: Search: LS based	Student	List of RMs	List of RMs based on LS
Student	7.0: Give Feedback	Student’s Feedback	RMs’ feedback
7.0: Give Feedback	Student Database	Feedback Information	Student’s feedback of RM
Admin	2.0: Update Profile	Student’s Detail	Student’s information
Admin	4.0: Update Student’s Learning Profile	Student’s Test Result	Student’s LS test result information
Admin	8.0: Insert RM	RM’s Detail	RM’s information
8.0: Insert RM	9.0: Extract Information	RM’s Content	RM’s content details
9.0: Extract Information	RM Database	Extraction Information	RM’s extraction detail
RM Database	10.0: Manage Report	RM’s Detail	RM’s information
Student Database	10.0: Manage Report	Student’s Detail	Student’s information
10.0: Manage Report	Admin	Reports	Reports of student and RM

Table 6.4: Function and Events

Function Name	DFD Process	Events
Register	1.0	Student registration
Update Profile	2.0	Student's profile updates
LS Test	3.0	LS test perform
Update Student's Learning Profile	4.0	LS profile updates
Search: Keyword based	5.0	List of RMs based on keyword
Search: LS based	6.0	List of RMs based on LS
Give Feedback	7.0	Student feedback
Insert RM	8.0	RM insertion
Extract Information	9.0	Information extracts from RM
Manage Report	10.0	Reports issues

User Roles (Table 6.5) is defined based on User Catalogue (presented in Table 6.1). User roles and function and events information are used to produce User Role/Function Matrix (Table 6.6).

Table 6.5: User Roles

User Role	Job Title	Activities Description
Student	Student Registered Student	Insert student details Answering VARK LS test Insert keywords and search for RM Give feedback
Admin	Admin	Manage students Manage RM Perform extraction analysis on RM

Table 6.6: User Role/Function Matrix

Function/User	Student	Admin
Register	x	
Update Profile	x	x
LS Test	x	x
Update Student's Learning Profile	x	x
Search: Keyword based	x	
Search: LS based	x	
Give Feedback	x	
Insert RM		x
Extract Information		x
Manage Report		x

6.2.4 Logical System Specification

The objective of the logical system specification is to provide a detailed specification of the processing and dialogue requirements for the proposed system. This module uses the information from Requirement Analysis and Requirement Specification to develop the system. In this stage, technical system option and requirement of design is produced.

Technical system option lists out the hardware and software used to develop the system. These were developed and published using free and open source platform and maintained using open source version control:

- System platform: Web MVC (Model View Control) Technology using ASP.NET MVC 3 using Razor Template Engine
- Database : MySQL 5
- Tools/IDE/Editor: Visual Studio Web Developer 2010 Express Edition
- Language: C#
- Web/Application Server: Apache Web Server + mod_mono (XSP)
- Middleware: Mono (Open source .NET Framework)
- Operating System: Linux
- PDF Library: PDF Library iTextSharp which is a complement of iText library is used to handle and manage PDF files.

Requirement of design provide detail specification that is used to develop the system such as system module and menu structure. System module provides flow for each sub-module in the system. This prototype has five modules as discussed in section 6.2. System Module for LSIST is shown in Figure 6.6 below.

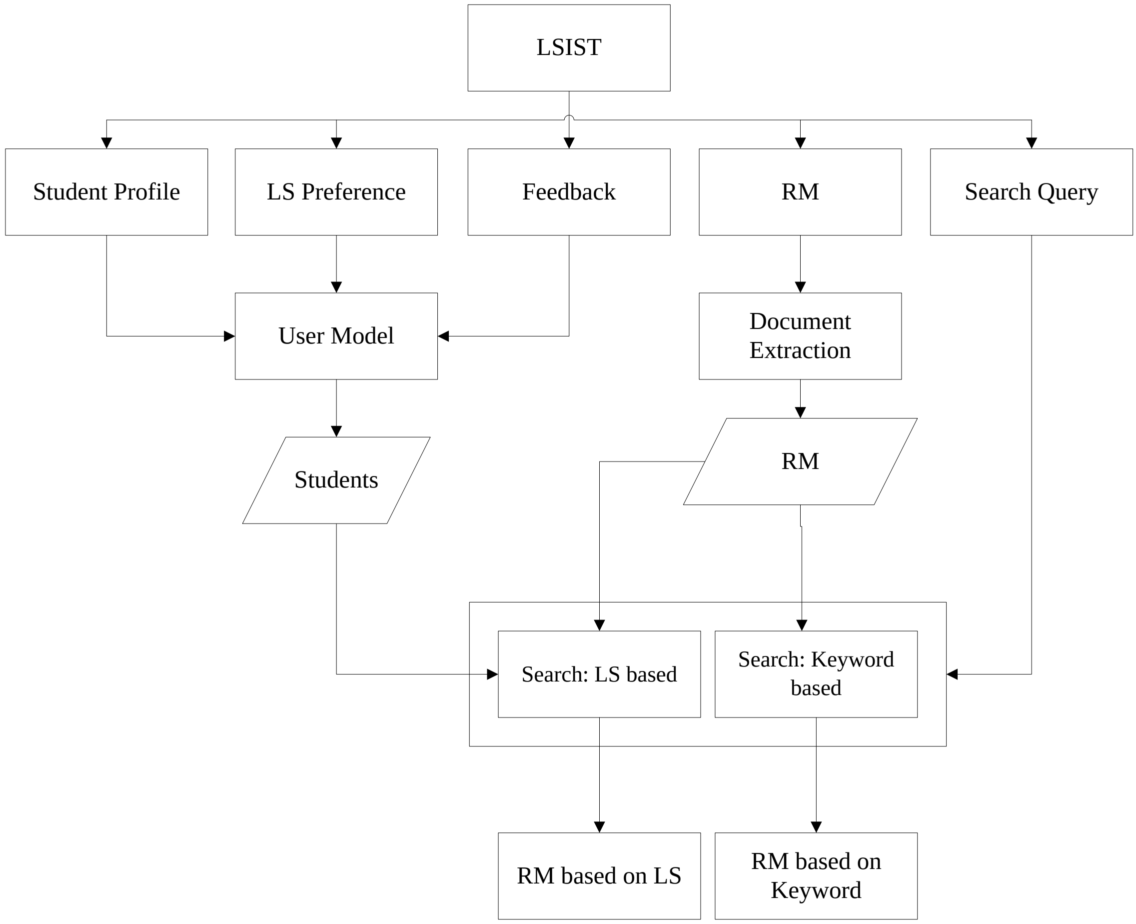


Figure 6.6: System Module

Menu structure is also produced in this phase. A menu structure is developed for each User Role. The User/Function Matrix is used to identify all the functions initiated by User Roles. The menu structure for Students User Roles is shown in Figure 6.7 while the menu structure for Admin User Roles is shown in Figure 6.8.

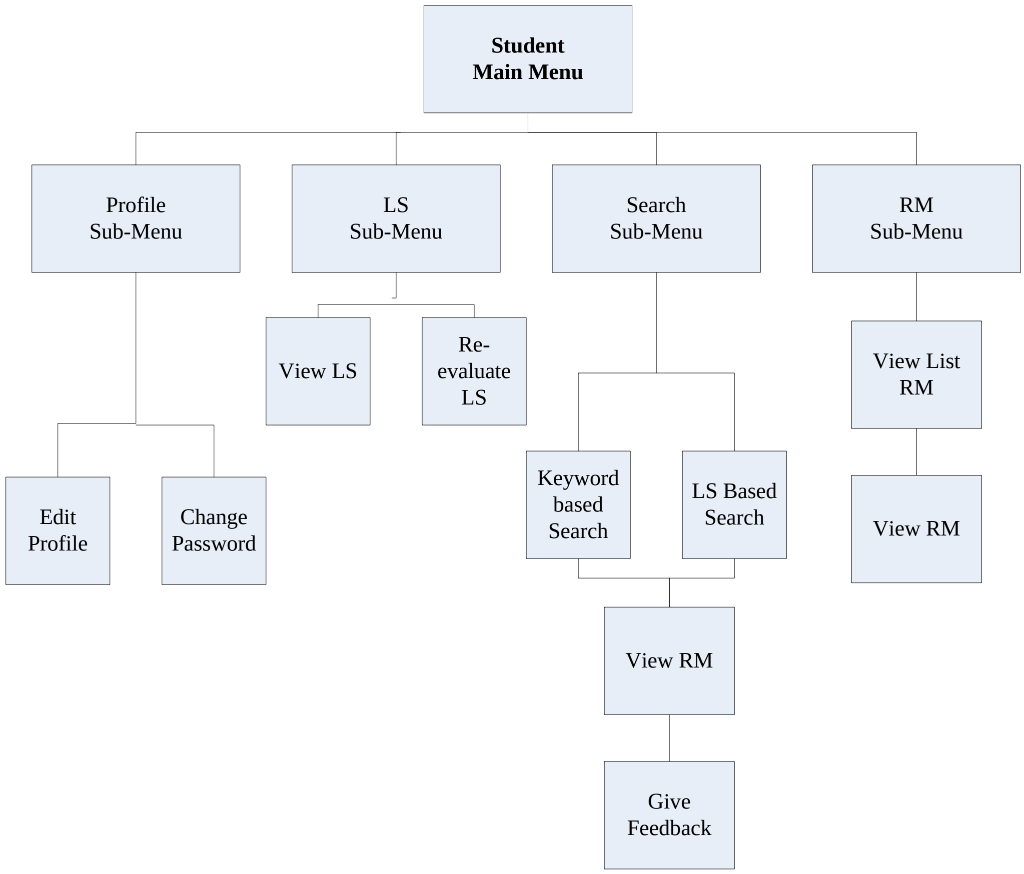


Figure 6.7: Student Main Menu

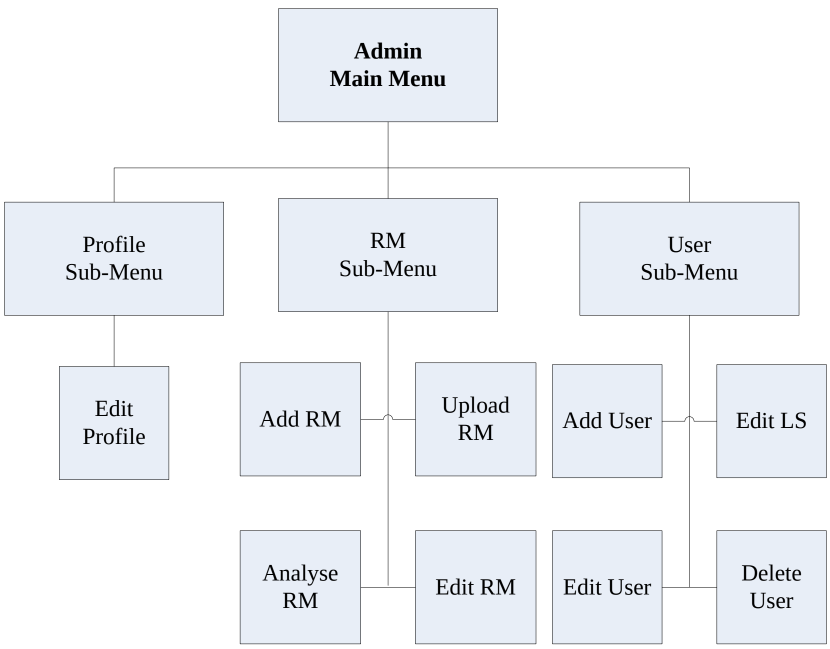


Figure 6.8: Admin Main Menu

6.2.5 Physical Design

Physical design is the final stage of SSADM when the logical design is converted to a design that fits the computer hardware and software selected. Physical design includes database design, interface and documentation. Figure 6.9 shows part of database.

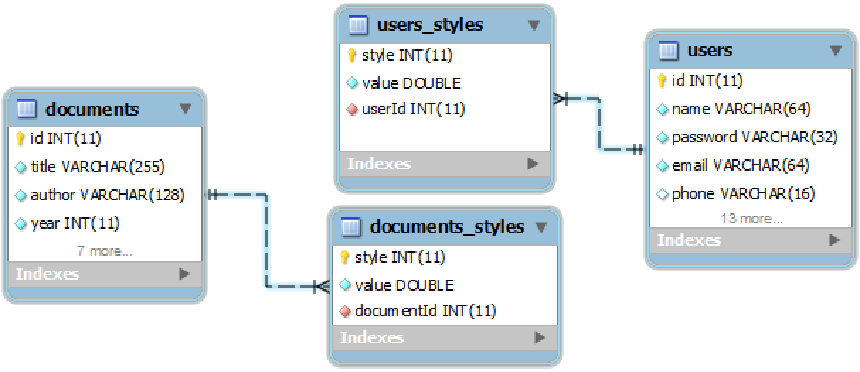


Figure 6.9: Sample of Database

Figure 6.10 – 6.19 are the LSIST interfaces for student.

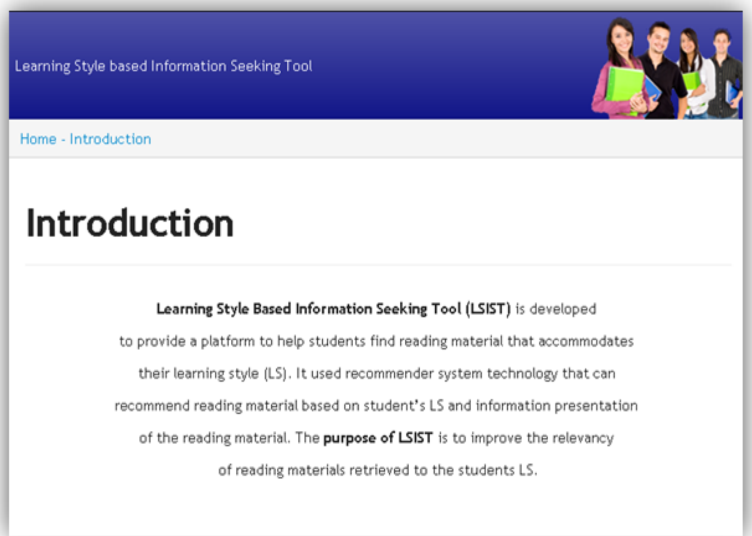


Figure 6.10: LSIST Homepage

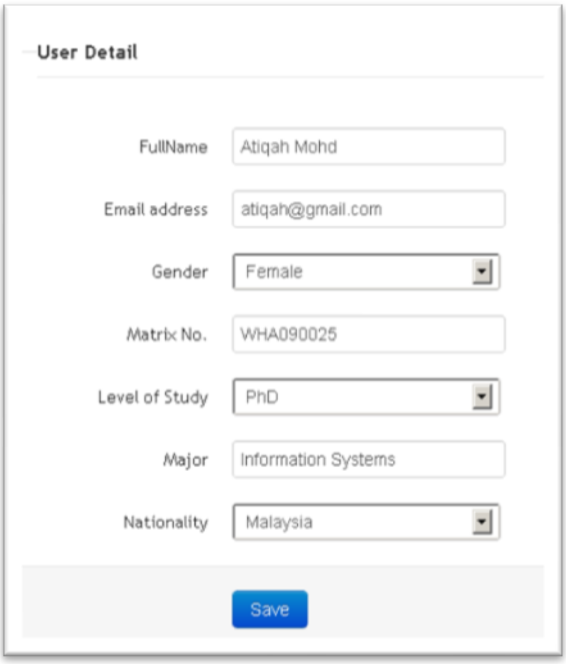


Figure 6.11: Students' Registration Form

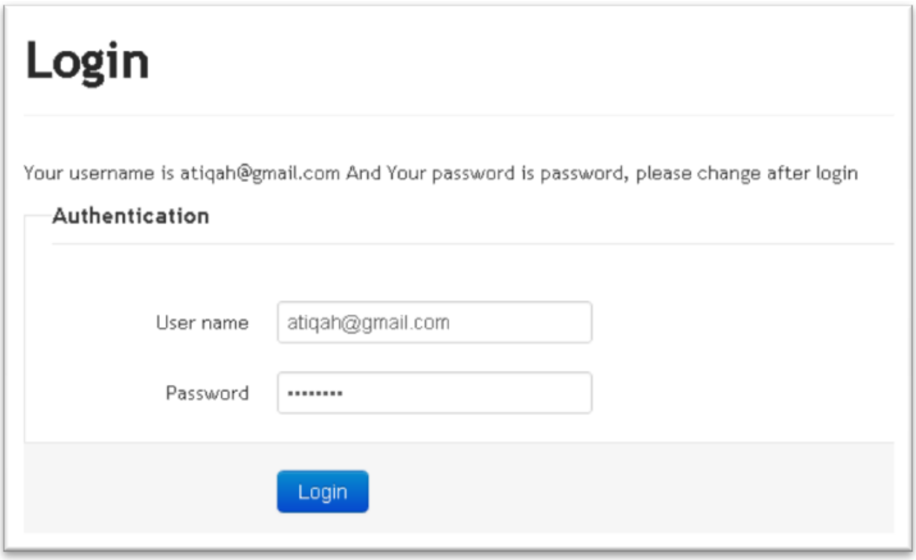


Figure 6.12: Log in Page for Student

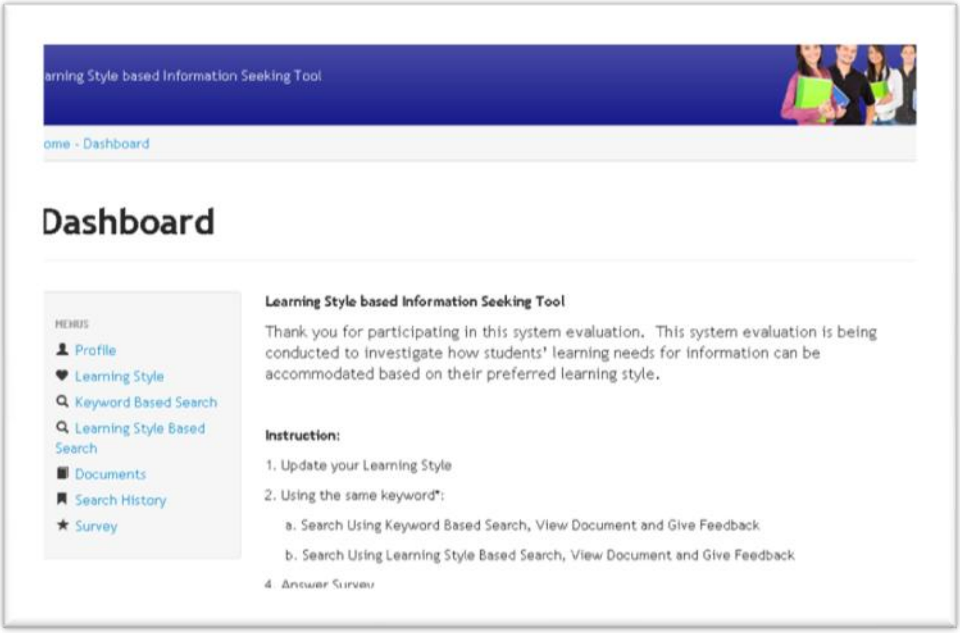


Figure 6.13: Student Homepage

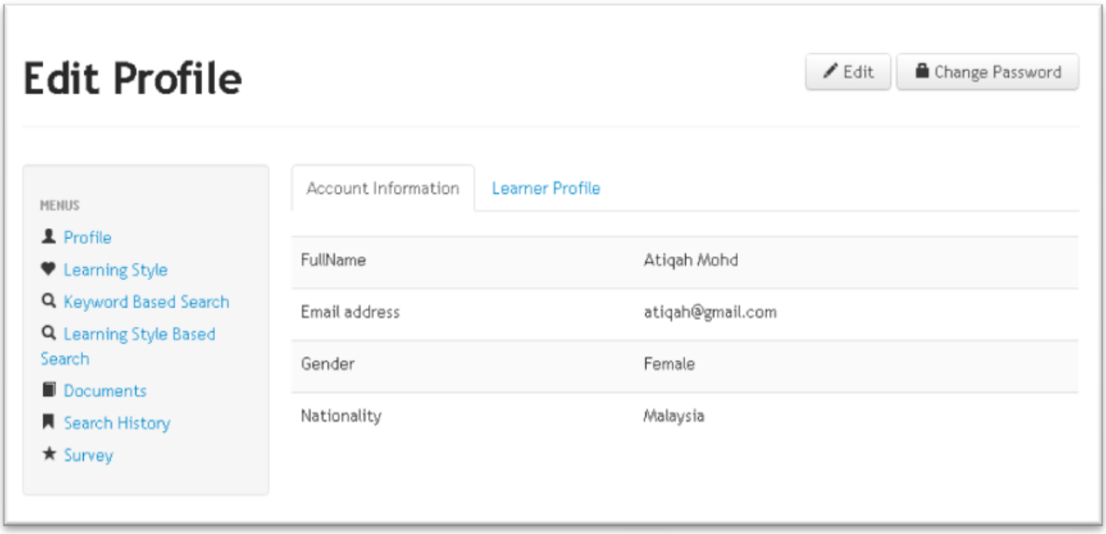


Figure 6.14: Edit Profile Page

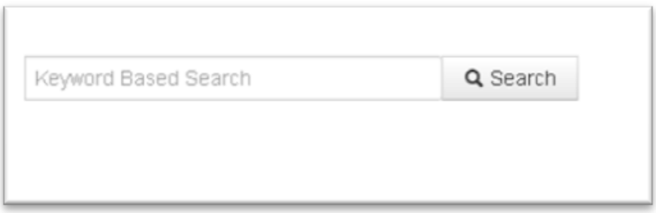


Figure 6.15: Keyword based Search Page

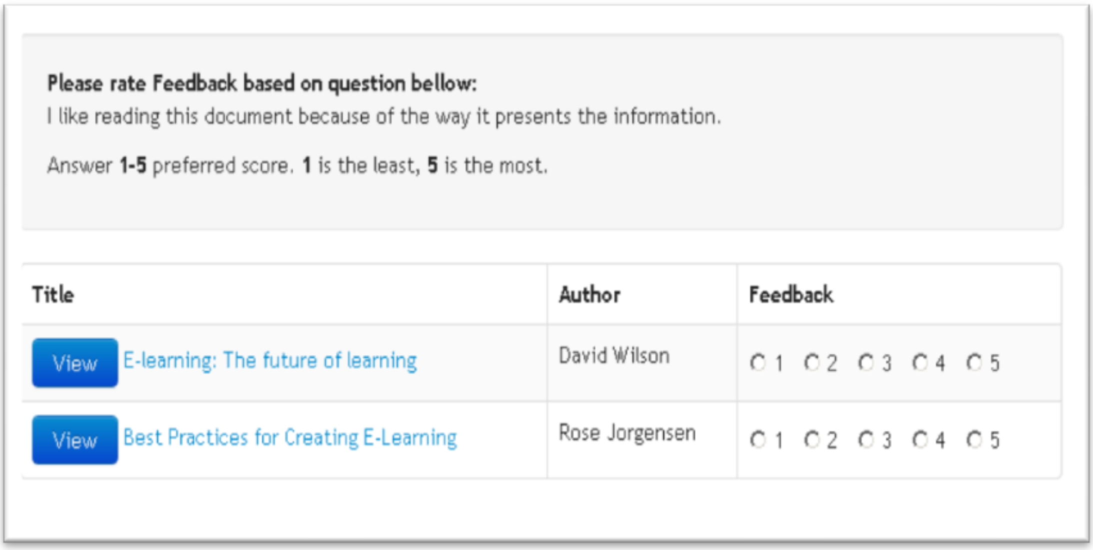


Figure 6.16: Search Results Page

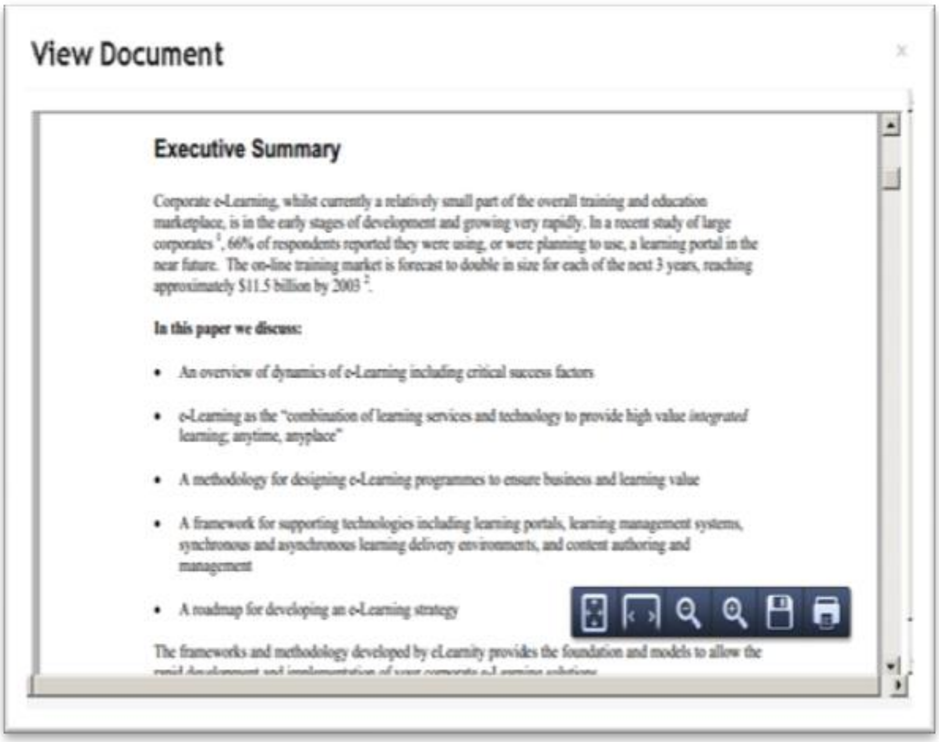


Figure 6.17: View RM Page

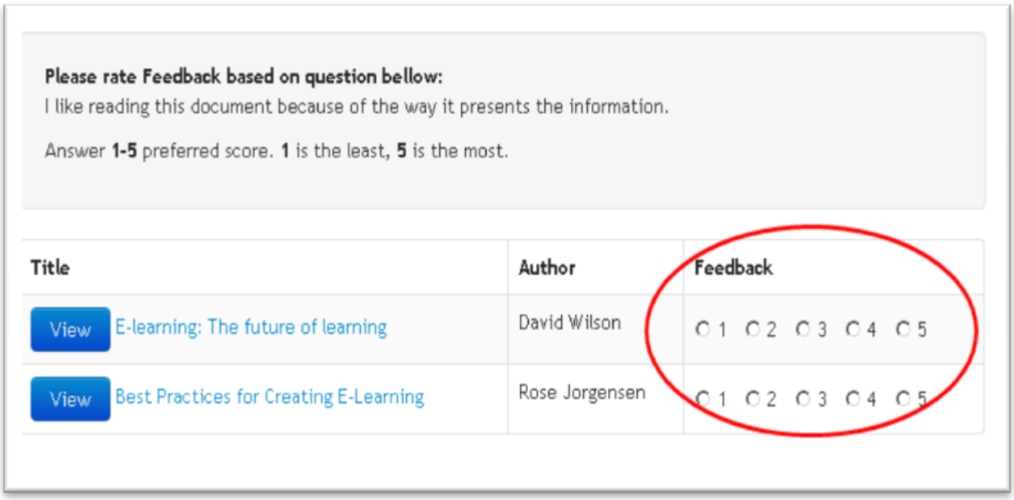


Figure 6.18: Give Feedback Page

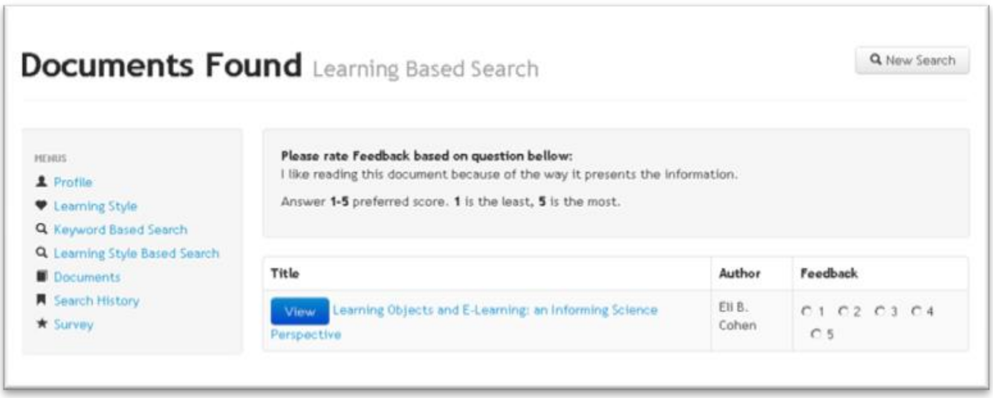


Figure 6.19: LS based Search Page

Figure 6.20 – 6.26 are the interfaces for admin.

Figure 6.20: Log In Page for Admin

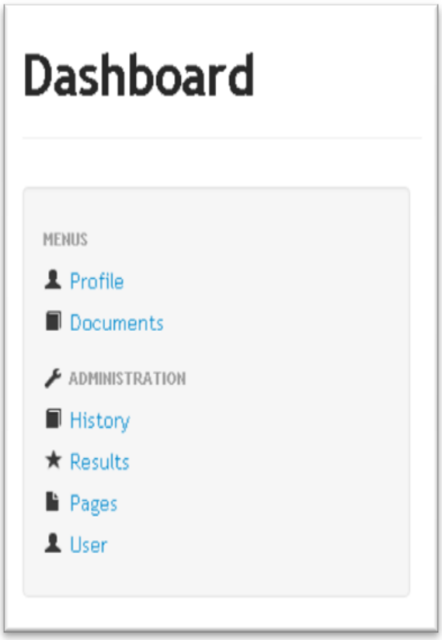
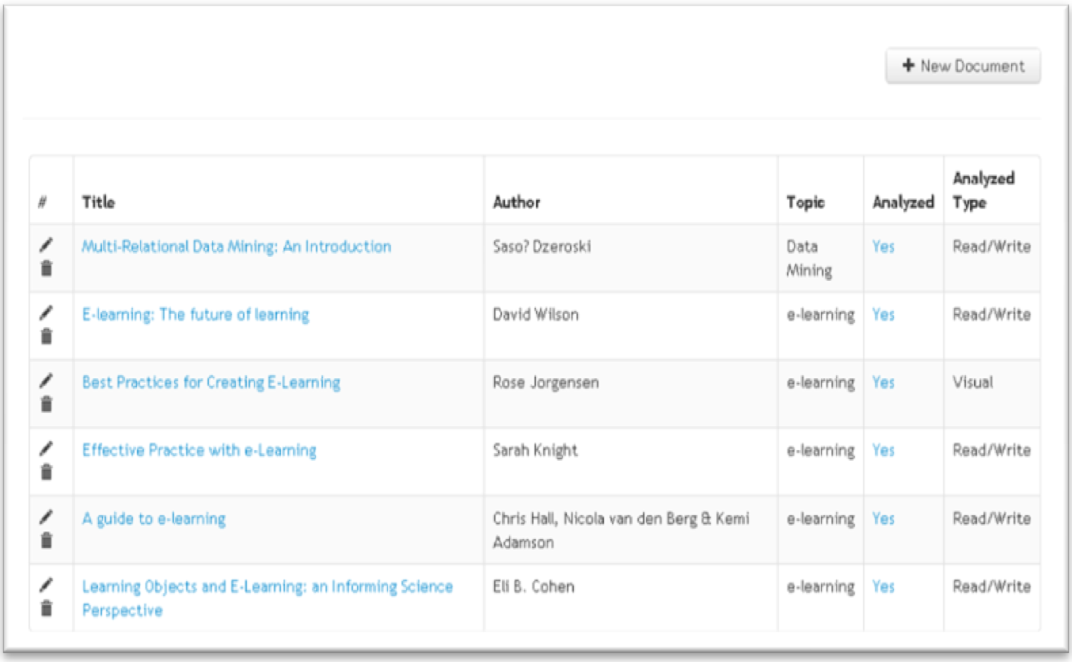


Figure 6.21: Admin Dashboard



+ New Document

#	Title	Author	Topic	Analyzed	Analyzed Type
	Multi-Relational Data Mining: An Introduction	Saso? Dzeroski	Data Mining	Yes	Read/Write
	E-learning: The future of learning	David Wilson	e-learning	Yes	Read/Write
	Best Practices for Creating E-Learning	Rose Jorgensen	e-learning	Yes	Visual
	Effective Practice with e-Learning	Sarah Knight	e-learning	Yes	Read/Write
	A guide to e-learning	Chris Hall, Nicola van den Berg & Kemi Adamson	e-learning	Yes	Read/Write
	Learning Objects and E-Learning: an Informing Science Perspective	Eli B. Cohen	e-learning	Yes	Read/Write

Figure 6.22: RM Page

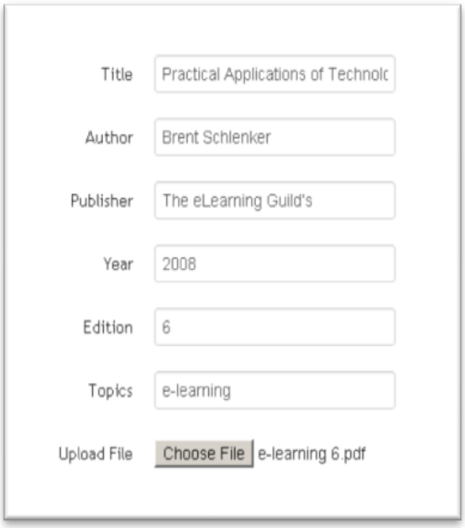
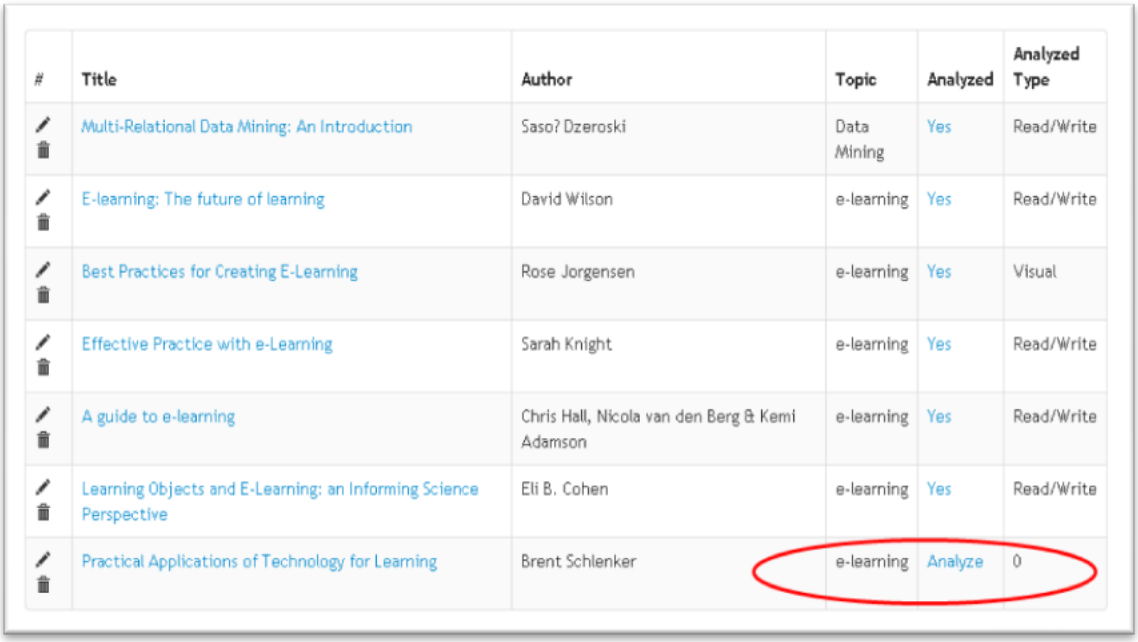


Figure 6.23: New RM Page



#	Title	Author	Topic	Analyzed	Analyzed Type
	Multi-Relational Data Mining: An Introduction	Saso? Dzeroski	Data Mining	Yes	Read/Write
	E-learning: The future of learning	David Wilson	e-learning	Yes	Read/Write
	Best Practices for Creating E-Learning	Rose Jorgensen	e-learning	Yes	Visual
	Effective Practice with e-Learning	Sarah Knight	e-learning	Yes	Read/Write
	A guide to e-learning	Chris Hall, Nicola van den Berg & Kemi Adamson	e-learning	Yes	Read/Write
	Learning Objects and E-Learning: an Informing Science Perspective	Eli B. Cohen	e-learning	Yes	Read/Write
	Practical Applications of Technology for Learning	Brent Schlenker	e-learning	Analyze	0

Figure 6.24: List of RM Page

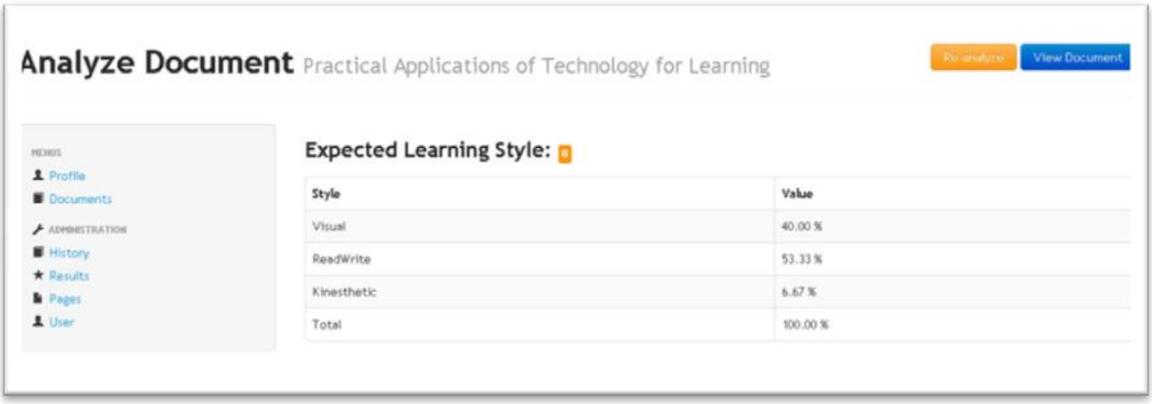


Figure 6.25: Results of Information Extraction

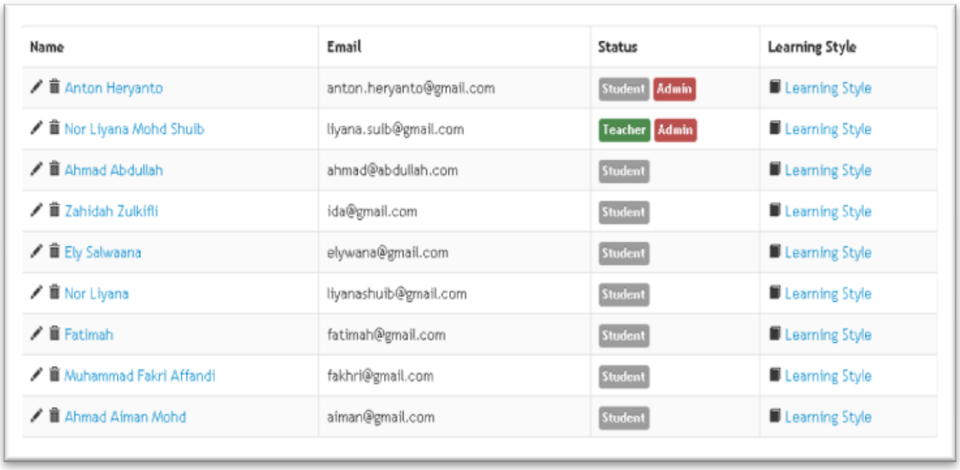


Figure 6.26: Manage User Page

6.3 System Implementation

System implementation includes hardware and software, data sets, and system testing. Hardware and software are listed in section 6.2.4. Data set and system testing are discussed below.

6.3.1 Data Set

Data used as input for this tool are from RM or documents of the PDF type. The data set consists of 77 PDF documents with various types of elements (Refer to Appendix C). It is in the form of journals, articles and eBook. The data set retrieved on top of Google search engine according to the steps below:

1. Search by keyword for RMs in PDF formats using Google search engine
2. RMs were filtered based on the variety of primitive elements presented in the RMs.

6.3.2 Algorithm Testing

To evaluate the effectiveness of *Features Extraction* algorithm, recall and precision measurement are used. Recall and precision is the most frequent measure used when referring to information extraction technique (Kaiser & Miksch, 2005). Recall and precision is measured based on expected LS identifiers extracted by feature extraction algorithm. For feature extraction, each page has only three feature vectors values which are *Visual*, *Read/write* and *Kinesthetic*. Features extraction calculation conducted are based on LS identifiers detected in each page of RM. Recall and precision for *Feature Extraction* algorithm is defined by:

$$\text{Recall} = \frac{F}{E}$$

$$\text{Precision} = \frac{F}{R}$$

Where;

F = Number of feature vectors that are found for the whole RM by *Feature Extraction* algorithm

E = Number of feature vectors expected from RM extracted manually by expert

R = Number of page in RM

Expert in Image Analysis examined sample of RM in database and extract information of *E* and *R* manually as shown in Appendix D. Data used for testing are shown in Table 6.7.

Table 6.7: Data for Feature Extraction Algorithm Testing

Feature Vector	***R***	***E***	***F***
Visual	97	37	32
Read/write	97	97	97
Kinesthetic	97	67	67

By using the above Recall and Precision formula and data in Table 5.8, results for Precision and Recall are shown in Table 5.9. Feature extraction results for all RM are presented in Appendix E.

Table 6.8: Precision and Recall for Feature Extraction Algorithm

Feature Vector	**Recall**	**Precision**
Visual	86.49%	32.99%
Read/write	100.00%	100.00%
Kinesthetic	100.00%	69.07%

The results show that Recall for each feature vector is high. Precision for *Read/write* and *Kinesthetic* feature vector is acceptable. However, the precision for the Visual feature vector is low with only 32.99%. This is due to the Feature Extraction Algorithm limitation which cannot extract vector image, symbol and image size. Hence some images in the RM were not identified.

6.3.3 System Testing

System testing is executed after system development has been completed. The process consists of two levels, namely:

a. Complete system testing

Complete system testing involves overall system function. During this test, user enters either sample data or real data, execute queries and view reports. Table 6.9 is an example of black box testing conducted to ensure the overall system functionality. Black box testing is a testing strategy which assumes that the output provided from the system is accurate based on specific input. This test needs to ensure the overall system function can be used accurately.

Table 6.9: Examples of Black Box Testing for Admin

Data Testing	Expected Result	Actual Results
Admin insert new student. Inputs consist of Student Profile and LS Preference	The addition of new student in Student database	The addition of new student in Student database
Admin insert new RM. Inputs consist of RM Information and Upload RM	The addition of new RM in RM database	The addition of new student in RM database
Admin run RM extraction analysis on the RM	Information of LS attribute is extracted from RM and inserted in RM database	Information of LS attribute is extracted from RM and inserted in RM database

b. Acceptance testing

This testing process begins after the system developed is defined as a complete package. Testing was conducted with the aim of ensuring that systems meet user requirements as defined in the specification of user requirements. In this test, the user is given the opportunity to use and test the system, and share their comments on the performance of the system. This testing process is discussed in the next chapter.

6.4 Summary

This chapter discussed the design and development of LSIST prototype. The architecture was first presented followed by system design using SSADM and system implementation. Both *Keyword based Search* and *LS based Search* components were included in the prototype for evaluation purposes. The evaluation of the prototype and its results will be presented in the next chapter.

CHAPTER 7

PROTOTYPE EVALUATION AND RESULTS

In the previous chapters, we have described LSIST design and development. To evaluate the prototype for LSIST, an experiment was carried out and a survey was conducted to determine the relevancy of the RMs to the students' LS. Data from the experiment and survey are analyzed.

7.1 Results

The experimental results presented and discussed below are based on the output of the experiment and survey. The combined experimental and survey results are given in Appendix G. Each respondent is labeled as R1, R2 and so on in SPSS. Table 7.1 presented LS preference based on Gender.

Table 7.1: LS Preference Based on Gender

LS preference	Gender	N (50)
Visual	Male	5
	Female	5
Aural	Male	3
	Female	7
Read/write	Male	4
	Female	6
Kinesthetic	Male	5
	Female	5
Multimodal	Male	3
	Female	7

7.1.1 Output of Experiment

A paired-samples t-test was conducted to compare respondents' feedback of Pretest and Posttest data. Table 7.2 shows the summary for Pretest and Posttest *Mean (M)* and *Standard Deviation (SD)*. There was a significant difference in the scores for Pretest (M=3.13, SD=0.67) and Posttest (M=4.14, SD=0.70) conditions; t (49) = -7.556, p = 0.000. These results suggest that LS does have an effect in finding suitable RM.

Table 7.2: Paired Samples Statistics

	Min	***Max***	***M***	***SD***
Pretest	2	4	3.13	0.67
Posttest	3	5	4.14	0.70

Mean for each group was compared to see the results between groups of LS preferences in relation to their satisfaction with the RM retrieved as shown in Table 7.3. The test revealed that all groups have better results on Posttest.

Table 7.3: Pretest and Posttest *Mean* for Each Group

LS Preference (N=50)	**Pretest**		**Posttest**	
	M	***SD***	***M***	***SD***
Visual	3.17	0.77	4.04	0.74
Aural	3.20	0.79	4.27	0.80
Read/write	2.97	0.50	3.83	0.76
Kinesthetic	3.20	0.63	4.10	0.57
Multimodal	3.10	0.74	4.47	0.57
Average	**3.13**	**0.67**	**4.14**	**0.70**

It indicates that the prototype do provide better RM to respondents. This is shown in Paired-samples t-tests that were conducted to all groups. Paired-samples t-tests were conducted to compare the feedback of Pretest and Posttest result for each group.

- *Visual*

  There was a significant difference in the scores for Pretest (M = 3.17, SD = 0.77) and Posttest (M = 4.04, SD = 0.74) results for *Visual* learner; t (9) = - 2.72, p = 0.02.

- *Aural*

  There was a significant difference in the scores for *Pretest* (M = 3.20, SD = 0.79) and Posttest (M = 4.27, SD = 0.80) results for *Aural* learner; t (9) = - 3.06, p = 0.01.

- *Read/write*

  There was a significant difference in the scores for Pretest (M = 2.97, SD = 0.50) and Posttest (M = 3.83, SD = 0.76) results for *Read/write* learner; t (9) = - 3.40, p = 0.01.

- *Kinesthetic*

  There was a significant difference in the scores for Pretest (M = 3.20, SD = 0.63) and Posttest (M = 4.10, SD = 0.57) results for *Kinesthetic* learner; t (9) = -3.86, p = 0.00.

- *Multimodal*

  There was a significant difference in the scores for Pretest (M = 3.10, SD = 0.74) and Posttest (M = 4.47, SD = 0.57) results for *Multimodal* learner; t (9) = -3.86, p = 0.00.

7.1.2 Survey Results

This survey was conducted to evaluate whether the proposed tool was able to retrieve RMs that are relevant to the students' LS. Figure 7.1 shows the results from all respondents.

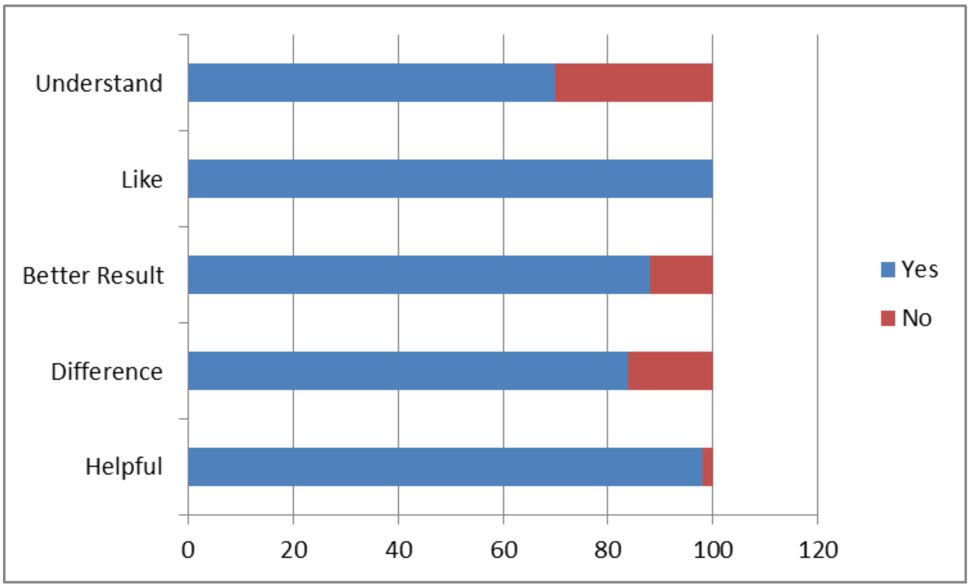


Figure 7.1: Results from All Respondents as a Percentage

The results show that:

- 70% of the respondents agreed that RMs retrieved by *LS based Search* were easier to understand.
- All fifty respondents said that they liked the way information was presented in the RMs suggested by the LS based Search.
- 88% of the respondents agreed that *LS based Search* gave better RM results than the *Keyword based Search*.
- 84% agreed that they can see the difference between *LS based Search* and *Keyword based Search*.

- 98% of the respondents agreed that this prototype was helpful in providing suitable RMs.

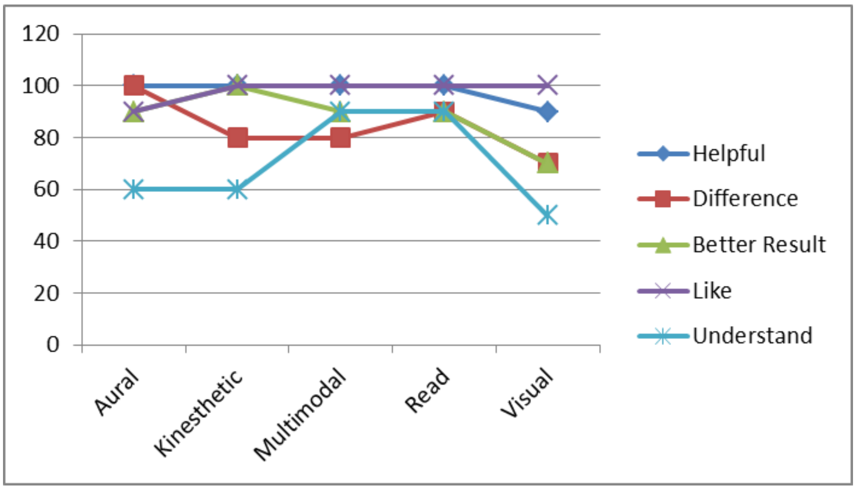


Figure 7.2: Results by Group as a Percentage

Figure 7.2 show the findings based on groups. The results show that:

- All groups agreed that this prototype was helpful in providing suitable RMs.
- All groups can see the difference between RM retrieved by *LS based Search* and *Keyword based Search*.
- Read, Aural, Kinesthetic and Multimodal group agreed that they retrieved better result from *LS based Search.* However, only 70% of Visual group agreed with that statement.
- The results from the survey indicate that all groups liked the RMs retrieved by the *LS based Search* more than the RMs retrieved by *Keyword based Search*.
- 90% of Read and Multimodal groups agreed that RMs retrieved by *LS based Search* were easier to understand. However, only 60% of Aural and Kinesthetic groups, and 50% of the Visual group agreed with that statement.

Below are comments from respondents. The comments show that the prototype is able to retrieve relevant RMs based on LS.

- *"It helps."* (R2)
- *"It somehow suits me better, I am more interested in the article based on LS at first glance, rather than keyword based searching."* (R17)
- *"Suit me."* (R27)
- *"Easy to understand."* (R28)
- *"LS based search was better than the keyword method. It is because I liked how the reading material was presented. It seemed short and concise."* (R30)
- *"Document that I like was put in first ranking."* (R37)
- *"Very nicely laid out and can be very useful. Just needs more articles/books of course."* (R46)

7.2 Discussion

The discussion is based on results from the experiment and the survey.

7.2.1 Results from the Experiment

A paired-samples t-test was conducted to compare the score of Pretest and Posttest. The significance value for Pretest and Posttest paired samples is 0.000 which is less than .05. Hence, we can conclude that there is a statistically significant difference between the score of Pretest and Posttest. Since our Paired Samples Statistics box revealed that the *Mean* number of Posttest was greater than the *Mean* for the Posttest, we can conclude that respondents like the RM retrieved by *LS based Search* more than RM retrieved by *Keyword based Search*.

Then, paired-sample t-tests were conducted on each group to compare the score of Pretest and Posttest for each group. The results show that there is a statistically significant difference between the score of Pretest and Posttest for all groups. These results suggest that LS really does have an effect in finding suitable RM.

7.2.2 Result from Survey

A simple survey was then conducted. The first part evaluated the respondents' opinion about the prototype. Most (98%) agreed that the prototype was helpful in providing suitable RM. They were then asked if they noticed any difference between the *LS based Search* and the *Keyword based Search*. Most (84%) agreed that there was a difference. Perhaps the most interesting trend with the score was that *Aural* learners had the highest percentage; even though their preference was not considered in the matching process. The results suggest that *Aural* learners could still retrieve suitable RM based on their alternative preference. Next, respondents were asked whether the *LS based Search* provided better RM than the *Keyword based Search*. The results revealed that the *Kinesthetic* group gave a higher result than the other groups; despite the fact that their preference was only extracted by keywords.

In the second part, respondents were asked about the RMs that they retrieved from LSIST. All (100%) showed that they liked the way information was presented in the RM suggested by the LS based Search. Next, they were asked if the RM retrieved from the *LS based Search* was easier to understand. Mixed results were achieved. The results revealed that *Read/write* and *Multimodal* groups gave a higher result than the other groups. The *Read/write* group's result was not a surprise, because the RM had the *Read/write* component. The *Multimodal* group was adaptable, since they had multiple preferences. The *Visual* group had the lowest result. This was due to the limitation of the Feature Extraction Algorithm. Vector images could not be identified and could only be extracted by keywords. However, the *Visual* group's results in Pre-test and Post-test did show that they could find RMs that suited them better.

7.3 Summary

The objective of this evaluation is to evaluate whether the prototype is able to retrieve RMs that are relevant to the students' LS. The results from the experiment show that *LS based Search* retrieved more suitable RM than *Keyword based Search*. The results from the survey indicate that all groups like the RM retrieved by *LS based Search* more than the RM retrieved by *Keyword based Search*.

These results suggest that LS does have an effect in finding suitable RM. Specifically, our results suggest that when respondents search RM based on LS, the possibility that they can find suitable RM increases. It shows that the prototype is able to find RMs that is more closely related to the students' LSs. Hence, we can conclude that the prototype is able to provide RM that is relevant to students' LS.

CHAPTER 8

CONCLUSION

In this research, we have investigated how students' learning needs for RM can be accommodated based on their preferred LSs. There are numerous studies devoted on using personalized IST in providing various learning material to students. However, there are no studies which specifically address the issue of providing suitable RM based on students' LS. Within the context of this study, each individual research question has been answered and discussed in Chapters 2, 3, 4, 5 and 6. This chapter discussed findings for this research followed by research contributions, its limitations and future work. From the onset of the study, the objectives and research questions (RQs) are determined as:

1. To investigate the need for an IST with LS consideration

   RQ 1: Are students having difficulties to find suitable RM for their learning needs?

   RQ 2: Is there any relation between RM and LS?

   RQ 3: Are students aware of their LS?

   RQ 4: Are there available ISTs that support students' LS?

2. To determine a method to categorize RMs based on LS

   RQ 5: How can primitive elements in RM be categorized based on LS?

3. To determine a method to map RM onto students' LS preference

   RQ 6: How can RM be mapped onto students' LS preference?

4. To develop and evaluate the LS based IST prototype

   RQ 7: Can the prototype perform better?

8.1 Research Findings

We revisit the RQs and objectives for this study. Findings for each RQ are discussed below.

RQ 1: Are students having difficulties to find suitable RM for their learning needs?

From the literature review and results of surveys discussed in section 4.2.1, clearly, there are difficulties faced by students in seeking information for their learning needs. They are:

- *Difficulties in finding suitable RM due to information overload*

  With so many RM available, students have difficulties to choose suitable RM for their learning needs.

- *Difficulties in finding suitable RM due to the different attribute in RM*

  Students have problems finding the most suitable RM due to the different attributes of materials, like the presentation of information (Yang & Wu, 2009).

RQ 2: Is there any relation between RM and LS?

From the previous study discussed in section 2.5.3 and results from LSIST evaluation discussed in section 7.2, it shows that there is a relation between RM and LS. Paired-sample t-tests were conducted on each LS preference groups to compare the result of Pretest (Keyword based Search) and Posttest (LS based Search). The results show that there is a statistically significant difference between the results of Pretest and Posttest for all groups. These results show that LS has an important role in the process of finding suitable RM. It can be concluded that there is a relation between RM and LS.

RQ 3: Are students aware of their LS?

From Figure 4.6, we can see that most respondents are unaware of their own LS preferences with 52.3%. This is in line with the findings from Rogers (2009) and Hassan et al. (2010). A comparison of Figure 4.8 and Figure 4.9 shows that the LS preference from respondents' perception is different from the actual LS preference identified from the VARK LS test. This supports the findings that students are unaware of their own LSs.

RQ 4: Are there available ISTs that support students' LS?

Four ISTs are compared to show the differences between several ISTs in terms of the capabilities and functionalities provided. The ISTs are Internet search engine, OPAC, online database and digital library. From the comparison done in Table 2.5, none of the existing ISTs has the function that can match RMs with students' LSs. There are a few personalized ISTs that include LS components but they only focus in providing different learning strategies. They are not meant to find suitable RMs that is aligned with students' LS as shown in Table 2.6.

RQ 5: How can primitive elements in RM be categorized based on LS?

Primitive elements in RM such as table, graph and diagram are identified and classified according to *Text* and *Non-Text*. Sub-categories of *Non-Text* categories can be categorized as *Photograph*, *Graphic*, *Semi Graphic* and *Diagram*. The classified primitive elements are then mapped onto LS preferences using identifiers from Learning Resource Type in LOM. These were discussed in section 5.3.

RQ 6: How can RM be mapped onto students' LS preference?

RM was mapped onto students' LS preference using k-NN classification method. The content of PDF was extracted and pre-processed before the process of feature extraction was executed. Feature vectors that have been retrieved from feature extraction process were calculated to get the LS value for RM using *Feature Extraction Algorithm*. Feature values for RM and student were then normalized to standardize the LS value. Then, feature vectors of RM and student were matched using k-NN in *Matching Algorithm*. These were discussed in section 5.4.

RQ 7: Can the prototype perform better?

Evaluation for the prototype was done using quasi-experiment design: Pretest/Posttest, Non-equivalent Multiple Group Design. LS based Search is compared with non LS based Search (Keyword based Search). The findings show that LS based Search can perform better than non LS based Search (see section 7.2.1). Survey was also used to get student feedback. The findings show that LSIST prototype can retrieve RM that matched students' LS as discussed in section 7.2.

Based on the answers for all RQs, research objectives are then discussed.

Research Objective 1: To investigate the need for an IST with LS consideration

Students faced difficulties in finding suitable RM for their learning need. One of the difficulties is that students have difficulties in finding suitable RM for their learning needs. The result of this investigation shows that there is a relation between RM and LS. Students need to know their LS to find suitable RM that accommodates their own LS. However, most students are unaware of their own LS preferences. Hence, they need IST that can help them find suitable RM based on their LS. The result indicates that none of the existing ISTs has the function that can match RMs with students' LSs. Hence, there is a need for an IST with LS consideration. These are discussed in RQ 1, RQ 2, RQ 3 and RQ 4.

Research Objective 2: To determine a method to categorize RMs based on LS

RM can be classified according to LS by using primitive elements such as text, diagram and table that contains in RM content. Primitive elements were identified and classified according to the form of primitive elements and mapped onto LS preferences using identifiers from LOM as discussed in RQ 5. Identifiers are then used to identify primitive elements in *Feature Extraction Algorithm*.

Research Objective 3: To determine a method to map RM onto students' LS preference

RM was mapped onto students' LS preference using k-NN classification method. The mapping was done using *Matching Algorithm* discussed in RQ 6. Only three LS preferences were considered in the matching processes: *Visual*, *Read/write* and *Kinesthetic* preference. However, the results from the evaluation as discussed in section 7.2.1 show that the prototype can also work with the other LS preferences such *Aural* and *Multimodal*.

Research Objective 4: To develop and evaluate the LS based IST prototype

The architecture for LS based IST (LSIST) was proposed in section 5.1. A prototype for LSIST was developed. It was designed and developed using SSADM as discussed in section 6.2. Evaluation for the prototype was done using quasi-experiment design and a survey. The findings show that LSIST prototype is able to find RMs that is more closely related to the students' LSs as discussed in RQ 7.

Based on the answer of all research questions and objectives, we can positively confirm that all research objectives have been successfully ascertained.

8.2 Key Findings

The findings from this research make several contributions to the current literature,

- Identification of students' difficulties in finding student
- Identification of students' unawareness of LS
- Identification of IST limitation
- LSIST Architecture
- Identification of VARK LS model for RM to be mapped to students' LS
- Identification and classification of primitive elements in RM according to LS
- Mapping of the classified primitive elements onto LS preferences
- Identification of identifiers for feature vectors using LOM
- *Feature Extraction Algorithm*, a novel algorithm that extracts and classifies RM onto LS preference.
- *Matching Algorithm*, a novel algorithm that match RM onto students' LS preference.
- LSIST prototype that retrieves RM based on students' LS.
- LSIST evaluation result based on quasi-experimental and a survey

8.3 Limitation

We identified a number of limitations that can be further studied for the research,

- Most of the survey items in the preliminary study and the evaluation used a categorical scale. This limited the capability of the data analysis.
- *Feature Extraction Algorithm* cannot extract vector images, symbols and size of image from PDF document. Vector images are combinations of arrow, line and word which represent the table, flow, chart and other object whilst symbols are mathematical equations and formula. Logical text is used to detect vector images and symbols. To identify and extract vector image and symbol as image in PDF file, knowledge in image analysis and object identification is needed. However, image analysis and object identification is beyond the focus of this study.
- Students' LS are identified using VARK LS test without considering their feedback in the evaluation process. Even though the feedback component already exists in the prototype, it is not taken into account in the matching process due to time constraint in developing the prototype.
- The sample size for the experiment in this research is limited to ten respondents for each type of LS preference to illustrate the capabilities of the prototype

8.4 Future Work

Taken from the limitations discussed in the previous sections:

- Due to *Feature Extraction Algorithm* limitation, techniques in image analysis and object identification can be explored to identify and extract *Non-Text* primitive elements. To categorize primitive elements in RM, it would be more comprehensive if all *Non-Text* primitive elements can be directly identified and extracted without just using logical text. *Feature Extraction Algorithm* can calculate the value of feature vector more precisely since the identifiers refer to the object itself and not just the logical text.

- Students' feedback can be added in the matching process. The proposed tool can be more personalized by incorporating an adaptive mechanism. The feedback can be used for adaptive purposes to retrieve RM based on students' LS and preferences. We believe by using students' feedback, the results can be made more precise and closer to students' preferred RM.

8.5 Conclusion

The primary focus of this research is to investigate how students' learning needs for RM can be accommodated based on their preferred LSs. The results of this investigation show that students have difficulties to understand the RM presented not aligned to their LS and to find suitable RM due to the limitation of existing LSTs. Thus, this research aims at developing an IST that incorporates LSs in its retrieval process.

This research has found that RM is not classified in a way that can be mapped to LS preference. One of the most significant findings to emerge from this research is the classification of RM onto LS preference using primitive elements contains in RM. Primitive elements in RM are classified and then mapped onto LS preference. VARK LS model is chosen because of its suitability with primitive elements categorization.

Two novel algorithms were embedded in this tool. *Feature Extraction algorithm* extracted and calculated identifiers in RM. *Matching algorithm* mapped RM onto student' LS preference using k-Nearest Neighbor classification model to match students to the suitable RM based on their LS preference. By developing this tool, students can retrieve RM based on their preferred LSs.

Finally, from the results of the experiments, we believe all the objectives set out at the beginning of the research have been achieved.

REFERENCES

Alkhasawneh, I., Mrayyan, M., Docherty, C., Alashram, S., & Yousef, H. (2008). Problem-based learning (PBL): assessing students' learning preferences using VARK. *Nurse education today*, *28*(5), 572–579.

Antoine, D., Collin, S., & Tombre, K. (1992). Analysis of Technical Document: The REDRAW System. In H. Baird, H. Bunke, & K. Yamamoto (Eds.), *Structured Document Image Analysis* (pp. 385 – 402). Springer Verlag.

Apriyani, T. I., & Hasibuan, M. S. (2008). Sistem E-Learning Dengan Pendekatan Gaya Belajar Vark. *Konferensi Nasional Sistem dan Informatika*, 19–23.

Ariyapala, P. G. (2002). *Use of the University of Malaya's Library OPAC by foreign postgraduate students* . University of Malaya. Retrieved from http://dspace.fsktm.um.edu.my/xmlui/handle/1812/185

Arms, W. Y. (2000). Digital Libraries. Cambridge, Mass., USA.

Ashworth, C., & Slater, L. (1992). *An Introduction to Ssadm Version 4*. United Kingdom: McGraw-Hill.

Atkin, C. (1973). Instrumental utilities and information seeking in new models for Mass Communication research. In P. Clarke (Ed.), *Reviews of Communication Research* (Vol. 2, pp. 205–242). Beverly Hills: Sage.

Atkins, H., Moore, D., Simon, S., & Dave, H. (2001). Learning Style Theory and Computer Mediated Communication. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 71–75). Tampere, Finland: Chesapeake, VA: AACE.

Aumann, Y., Feldman, R., Liberzon, Y., Rosenfeld, B., & Schler, J. (2006). Visual information extraction. *Knowledge and Information Systems*, *10*(1), 1–15. doi:10.1007/s10115-006-0014-x

Bachari, E. El, Abelwahed, E. H., & Adnani, M. El. (2011). E-Learning Personalization Based On D Ynamic Learners ' Preference. *International Journal of Computer Science & Information Technology*, *3*(3).

Baker, J., Sexton, A., & Sorge, V. (2008). Extracting Precise Data from PDF Documents for Mathematical Formula Recognition. *In 8th IAPR International Workshop on Document Analysis Systems*. Nara, Japan. Retrieved from ftp://ftp.cs.bham.ac.uk/pub/authors/V.Sorge/publications/2008/BaSeSo08das.pdf

Barrett, A. (2005). The Information-Seeking Habits of Graduate Student Researchers in the Humanities. *The Journal of Academic Librarianship*, *31*(4), 324–331.

Bayer, T., Franke, J., Kressel, U., Mandler, E., Oberl~der, M., & Schijrmann, J. (1992). Towards the Understanding of Printed Documents. In H. Baird, H. Bunke, & K.

Yamamoto (Eds.), *Structured Document Image Analysis*. New-York: Springer-Verlag.

Becker, N. J. (2003). Google in perspective: understanding and enhancing student search skills. *New Review of Academic Librarianship*, *9*(1), 84–100.

Bridle, J. S. (1990). Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. In F. F. Soulie & J.Herault (Eds.), *Neurocomputing: Algorithms, Architectures and Applications* (pp. 227–236). Berlin: Springer-Verlag.

Budhu, M. (2002). Interactive web-based learning using interactive multimedia simulations. *International Conference on Engineering Education (ICEE 2002)* (pp. 1–6). Manchester, UK: International Network for Engineering Education & Research.

Byron, M. (2004). *Satisficing and Maximizing: Moral Theorists on Practical Reason*. Cambridge: Cambridge University Press.

Case, D. O. (2007). *Looking for information: a survey of research on information seeking, needs, and behavior* (2nd ed.). Amsterdam: Academic Press.

Catalano, A. (2013). Patterns of graduate students' information seeking behavior: a meta-synthesis of the literature. *Journal of Documentation*, *69*(2).

Catalano, A. J. (2010). Evidence Based Library and Information Practice. *Evidence Based Library and Information Practice*, *5*(4), 21–38.

Cheng, Y. (2009). A Cognition Inference-Based Approach for Learning Object Recommendation in E-Learning. *International Symposium on Computer Network and Multimedia Technology* (pp. 1–4). Shanghai, China.

Chowdhury, G. G. (1999). *Introduction to modern information retrieval*. London: Library Association Publishing.

Chowdhury, G. G., & Chowdhury, S. (2003). *Introduction to Digital Libraries*. London: Facet Pub.

Christodoulakis, S., & Theodoridou, M. (1986). Multimedia document presentation, information extraction, and document formation in MINOS: a model and a system. *ACM Transactions on Office Information Systems*, *4*(4), 345–383. Retrieved from http://dl.acm.org/citation.cfm?id=9764

Commonwealth of Learning. (2005). *Creating learning materials for open and distance learning: a handbook for authors & instructional designers*. Canada: Commonwealth of Learning. Retrieved from http://dspace.col.org/handle/123456789/157

Connaway, L. S., Dickey, T. J., & Radford, M. L. (2011). If it is too inconvenient I'm not going after it: Convenience as a critical factor in information-seeking behaviors. *Library & Information Science Research*, *33*(3), 179–190.

Cooke, A. (2001). *A guide to finding quality information on the Internet: selection and evaluation strategies*. London: Library Association.

Cooper, D. R., & Schindler, P. S. (2006). *Business research methods*. (T. M.-H. series, Ed.). New York: McGraw-Hill Irwin.

Creswell, J.W. (2008). *Educational Research – Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Los Angeles: Pearson International Edition.

Creswell, John W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Los Angeles: Sage Publication, Inc.

Curry, L. (1983). *An organization of learning style theory and constructs. , 235, 185.*

Danielsson, P.-E. (1980). Euclidean distance mapping. *Computer Graphics and image processing*, *14*(3), 227–248. doi:10.1016/0146-664X(80)90054-4

Déjean, H., & Meunier, J. (2006). A system for converting PDF documents into structured XML format. *Document Analysis Systems VII*, 1–12. Retrieved from http://www.springerlink.com/index/yh44743411772022.pdf

DeJong, G. (1982). An overview of the FRUMP system. In W. G. Lehnert & M. H. Ringle (Eds.), *Strategies for natural language processing* (pp. 149–176). Hillsdale: Erlbaum.

Dobson, J. L. (2009). Learning style preferences and course performance in an undergraduate physiology class. *Advances in Physiology Education*, *33*, 308–314.

Dobson, J. L. (2010). A comparison between learning style preferences and sex, status, and course performance. *Advances in physiology education*, *34*(4), 197–204. doi:10.1152/advan.00078.2010

Dori, D., Doermann, D., Shin, C., Haralick, R., Phillips, I., Buchman, M., & Ross, D. (1997). The Representation of Document Structure: A Generic Object-Process. In P. S. P. Wang & H. Bunke (Eds.), *Handbook on Optical Character Recognition and Document Image Analysis*. World Scientific Publishing Company.

Dunn, R., & Dunn, K. (1978). *Teaching students through their individual learning styles*. Reston, VA: Reston.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, *45*, 171–212.

Felder, R. M. (1995). A Longitudinal Study of Engineering Student Performance and Retention. IV. Instructional Methods and Student Responses to Them. *J. Engr. Education*, *84*(4), 361–367.

Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education*, *78*(7), 674–681.

Fleming, N. (2010). The VARK Questionnaire. Retrieved January 1, 2010, from http://www.vark-learn.com/english/page.asp?p=questionnaire

Fleming, N., & Baume, D. (2006). Learning Styles Again: VARKing up the right tree!! *Educational Developments*, (7.4), 4–7.

Fleming, N. D. (1995). I'm different; not dumb Modes of presenting (V.A.R.K) in the tertiary classroom.

Fleming, N. D., & Mills, C. (1992a). Not another inventory, rather a catalyst for reflection. *To Improve the Academy*. Retrieved January 7, 2013, from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1245&context=podimproveacad

Fleming, N. D., & Mills, C. (1992b). VARK a guide to learning styles. Retrieved from http://www.vark-learn.com/English/index.asp

Fong, a. C. M., Hui, S. C., & Vu, H. L. (2002). Effective techniques for automatic extraction of Web publications. *Online Information Review*, *26*(1), 4–18. doi:10.1108/14684520210418347

Franzoni, A., & Assar, S. (2009). Student learning styles adaptation method based on teaching strategies and electronic media. *Educational Technology & Society*, *12*(4), 15–29. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4561832

Franzoni, A. L., Assar, S., Defude, B., & Rojas, J. (2008). Student Learning Styles Adaption Method Based on Teaching Strategies and Elecronic Media. *International Conference on Advanced Learning Technologies*.

Galvin, J. (2005). Alternative Strategies for Promoting Information Literacy. *The Journal of Academic Librarianship*, *31*(4), 352–357.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*. New York: Basic Books.

Geist, E. A., & King, M. (2008). Different, Not Better: Gender Differences in Mathematics Learning. *Journal of Instructional Psychology*, *35*(1), 43–52.

Godleski, E. (1984). Learning style compatibility of engineering students and faculty. *Proceedings Annual Frontiers in Education Conference*. Philadelphia.

Graf, S., Viola, S. R., Leo, T., & Kinshuk. (2007). In-Depth Analysis of the Felder-Silverman Learning Style Dimensions. *Journal of Research on Technology in Education*, *40*(1).

Grasha, A. & Reichmann, S. W. (1974). A Rational Approach to Developing and Assessing the Construct Validity of a Student Learning Style Scale Instrument. *Journal of Psychology*, *87*, 213– 223.

Gregorc, A. F. (1982). *An Adult's Guide to Style*. Maynard, Massachusetts: Gabriel Systems, Inc.

Gregorc, A. F. (1985). *Inside styles, beyond the basics*. Columbia, CT: Gregorc Associates.

Han, J., & Kamber, M. (2000). *Data mining: concepts and techniques*. San Francisco: Morgan Kaufmann Publishers Inc.

Hassan, R. (2009). How Learners Respond to Computer Based Learning Material Based on Modality Learning Style? *Teaching And Learning Open Forum 2009*. Curtin University of Technology Sarawak.

Hassan, R., Hashim, M. N., & Hussein, A. (2010). How Do Learners Respond To CBL Material Based On Modality Learning Style? *Journal of Technical Education and Training*, *2*(1). Retrieved from http://penerbit.uthm.edu.my/ojs/index.php/JTET/article/view/290

Hassan, T., & Baumgartner, R. (2005). Intelligent Text Extraction from PDF Documents. *International Conference on Computational Intelligence for Modelling Control and Automation* (pp. 2–6).

Honey, P., & Mumford, A. (1992). *The Manual of Learning Styles* (3rd ed.). Maidenhead: Peter Honey Publications.

Howles, L. (2006). Learning Styles: What the Research Says and How to Apply it to Designing E-Learning.

Huck, G., Fankhauser, P., Aberer, K., & Neuhold, E. (1998). Jedi: Extracting and Synthesizing Information from the Web. *Third International Conference of Cooperative Information Systems* (p. 32).

IEEE. (2002). Draft standard for learning object metadata. *Accessed July*. Retrieved November 14, 2012, from http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

Ismail, M. A., Kareem, S. A., & Nasaruddin, F. H. (2009). Assisting Novice Researchers in Utilizing the Web as a Platform for Research: Semantic Approach. *WSEAS Transactions On Information Science & Applications*, *2*(6).

Iverson, D. (2005). Information Literacy: what is it? And, why would I want to include it in my class? . Retrieved from http://www.sheridan.edu/library/services/Info literacy Jan 2005.ppt

James, S., D'Amore, A., & Thomas, T. (2011). Learning preferences of first year nursing and midwifery students: utilising VARK. *Nurse education today*, *31*(4), 417–23. doi:10.1016/j.nedt.2010.08.008

Julien, H., & Barker, S. (2009). How high-school students find and evaluate scientific information: A basis for information literacy skills development. *Library & Information Science Research*, *31*, 12–17.

Kaiser, K., & Miksch, S. (2005). Information Extraction: A Survey. Austria, Europe.

Keefe, J. W. (1989). *Learning style profile handbook: Accommodating perceptual, study and instructional preferences* (2nd ed.). Reston, VA: National Association of Secondary School Principals.

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook* (3rd ed.). Englewood Cliffs: Prentice Hall.

Klasnja-Milicevic, A., Vesin, B., Ivanovic, M., Budimac, Z., Klašnja-Milićević, A., & Ivanović, M. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education*, *56*(3), 885~899. doi:10.1016/j.compedu.2010.11.001

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, New Jersey: Prentice Hall.

Kolb, D. A. (1999). *The Kolb Learning Style Inventory* (3rd ed.). Boston: Hay Group.

Korobili, S., Malliari, A., & Zapounidou, S. (2011). Factors that Influence Information-Seeking Behavior: The Case of Greek Graduate Students. *The Journal of Academic Librarianship*, *37*(2), 155–165. doi:10.1016/j.acalib.2011.02.008

Kuhlthau, Carol C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. *Journal Of The American Society For Information Science*, *42*(5), 361–371.

Kuhlthau, Carol Collier. (1997). Learning in digital libraries: An information search process approach. *Library Trends*, *45*(4), 708–724.

Lancaster, F. W. (1968). *Information retrieval systems: characteristics, testing, and evaluation* (1st ed.). New York: John Wiley & Sons.

Lancaster, F. W. (1979). *Information retrieval systems: characteristics, testing, and evaluation* (2nd ed.). New York: John Wiley & Sons Inc.

Leite, W. L., Svinicki, M., & Shi, Y. (2009). Attempted Validation of the Scores of the VARK: Learning Styles Inventory With Multitrait-Multimethod Confirmatory Factor Analysis Models. *Educational and Psychological Measurement*, *70*(2), 323–339. doi:10.1177/0013164409344507

Li, J., & Gray, R. M. (2000). Context-based multiscale classification of document images using wavelet coefficient distributions. *IEEE Transactions on Image Processing*, *9*(9), 1604–1616. doi:10.1109/83.862641

Limberg, L. (1999). Experiencing information seeking and learning: A study of the interaction between two phenomena. *Information Research*, *5*(1). Retrieved from http://informationr.net/ir/5-1/paper68.html

Liu, Z., & Yang, Z. Y. L. (2004). Factors Influencing Distance-Education Graduate Students' Use of Information Sources: A User Study. *The Journal of Academic Librarianship*, *30*(1), 24–35.

Lu, X., Kataria, S., Brouwer, W. J., Wang, J. Z., Mitra, P., & Giles, C. L. (2009). Automated analysis of images in documents for intelligent document search. *International Journal on Document Analysis and Recognition (IJDAR)*, *12*(2), 65–81. doi:10.1007/s10032-009-0081-0

Lu, X., Mitra, P., Wang, J., & Giles, C. (2006). Automatic categorization of figures in scientific documents. *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries*. Retrieved from http://dl.acm.org/citation.cfm?id=1141778

Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge, Eng.: Cambridge University Press.

Marchionini, G. (1998). Digital Library Research and Development. *Encyclopedia of Library and Information Science*, *63*(26), 259–279.

McCarthy, B. (1987). *The 4MAT System: Teaching to learning styles with right/left mode techniques*. Barrington, IL: About Learning Inc.

Nagy, G., Seth, S., & Viswanathan, M. (1997). DIA, OCR, and the WWW. In P. H. Bunke & S. P. Wang (Eds.), *Handbook of Character Recognition and Document Image Analysis* (pp. 729–754). World Scientific Publishing Company.

Omar, I. I., & Mansor, Y. (2005). *Panduan mencari maklumat* (p. 155). Bentong: PTS Publications.

Oxford. (2005). *Oxford Dictionary of English, Revised Edition*. (C. Soanes & A. Stevenson, Eds.). Oxford University Press. Retrieved from http://oxforddictionaries.com/

Pen˜a, C. I., Marzo, J. L., & Rosa, J. L. (2002). Intelligent agents in a teaching and learning environment on the Web. *ICALT 2002*. Kazan, Russia: IEEE.

Peterson, E. R. (2009). Researching the psychology of cognitive style and learning style: Is there really a future? *Learning and Individual Differences*, *19*(4), 518–523.

Pitale, S., & Sharma, T. (2011). Information Extraction Tools for Portable Document Format. *International Journal of Computer Technology Application*, *2*(6), 2047–2051. Retrieved from http://www.doaj.org/doaj?func=abstract&id=894537

Raghavan, V., Bollmann, P., & Jung, G. S. (1989). A critical investigation of recall and precision as measures of retrieval system performance. *ACM Transactions on Information Systems*, *7*(3), 205–229. doi:10.1145/65943.65945

Rasmussen, K. L. (1998). Hypermedia and learning styles: Can performance be influenced? *Journal of Multimedia and Hypermedia*, *7*(4).

Rea, L., & Parker, R. A. (2005). *Designing and Conducting Survey Research: A Comprehensive Guide* (3rd ed.). Jossey-Bass.

Rehak, D. R., & Mason, R. (2003). Keeping the learning in learning objects. In A. Littlejohn (Ed.), *Reusing online resources: a sustainable approach to e-Learning* (pp. 22–30). London: Kogan Page.

Reid, J. M. (1987). The Learning Style Preferences of ESL Students. *TESOL QUARTERLY*, *21*(1), 87–110.

Reiser, R. A., & Dempsey, J. V. (2007). *Trends and Issues in Instructional Design and Technology* (2nd ed.). Saddle River, NJ: Pearson Education.

Renzulli, J. S. (1986). The three ring conception of giftedness : A developmental model for creative productivity. In R. S. & J. Davidson (Ed.), *Conceptions of giftedness* (pp. 53–92). Cambridge, MA: Cambridge University Press.

Rogers, K. M. A. (2009). A preliminary investigation and analysis of student learning style preferences in further and higher education. *Journal of Further and Higher Education*, *33*(1), 13–21.

Saad, M. S. M., & Zainab, A. N. (2007). Information Search and Use of Computer Science and Information Technology Undergraduates. *Proceedings of the International Conference on Libraries, Information and Society, ICOLIS 2007*, 291–304.

Saad, M. S. M., & Zainab, A. N. (2009). An investigation of information seeking behaviour of Computer Science and Information Technology undergraduates: a qualitative approach. *Malaysian Journal of Library & Information Sciences*, *14*(3), 15–34.

Sadler-Smith, E. (1996). Learning styles: a holistic approach. *Journal of European Industrial Training*, *20*(7), 29–36.

Saiti, A., & Prokopiadou, G. (2008). Post-graduate students and learning environments: Users' perceptions regarding the choice of information sources. *The International Information & Library Review*, *40*(20), 94–103.

Sapsford, R. (2007). *Survey Research* (2nd ed.). SAGE Publications, Ltd. doi:10.4135/9780857024664

Savic, G., & Konjovic, Z. (2009). Learning Style Based Personalization of SCORM E-learning Courses. *7th International Symposium On Intelligent Systems And Informatics* (pp. 316–320). IEEE.

Schiaffino, S., Garcia, P., & Amandi, A. (2008). eTeacher: Providing personalized assistance to e-learning students. *Computers & Education*, *51*(4), 1744–1754. doi:10.1016/j.compedu.2008.05.008

Shao, M., & Futrelle, R. (2006). Recognition and classification of figures in PDF documents. *Proceedings of the 6th international conference on Graphics Recognition: ten Years Review and Future Perspectives*, 231–242. Retrieved from http://www.springerlink.com/index/L16G884328039G48.pdf

Silver, H. F., Strong, R. W., & Perini., M. J. (2000). *So each may learn: Integrating learning styles and multiple intelligences*. Alexandria, VA: Association for Supervision and Curriculum Development.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, *69*(1), 99–118.

Slater, J. A., Lujan, H. L., & Dicarlo, S. E. (2007). Does gender influence learning style preferences of first-year medical students? *Advances in Physiology Education*, *31*, 336–42.

Stash, N., Cristea, A., & Brau, P. D. (2004). Authoring of learning styles in adaptive hypermedia: problems and solutions. *The 13th International Conference on World Wide Web* (pp. 114–123).

Sternberg, R., & Grigorenko, E. (1997). Are cognitive styles still in style? *American psychologist*, *52*(7), 700–712. doi:10.1037//0003-066X.52.7.700

Sun, J., & Xie, Y. (2009). A Recommender System Based on Web Data Mining for Personalized E-Learning. *International Conference on Information Engineering and Computer Science* (pp. 1–4). ICIECS 2009.

Tombre, K., & Dori., D. (1997). Interpretation of Engineering Drawings. In H. Bunke & P. S. P. Wang (Eds.), *Handbook of Character Recognition and Document Image Analysis* (pp. 457–484). World Scientific Publishing Company.

Tomlinson, C. (2009). Differentiating instruction as a response to academically diverse student populations. In R. Marzano (Ed.), *On Excellence in Teaching* (pp. 247–268). Bloomington: Solution Tree.

Weaver, P. L., Lambrou, N., & Walkley, M. (1998). *Practical SSADM 4+*. London: Pitman.

Wiersma, W., & Jurs, S. G. (2009). *Research Methods in Education*. Pearson International Edition.

Wilson, T. D. (1999). Models in Information Behaviour Research. *Journal of Documentation*, *55*(3), 249–270.

Wilson, T. D. (2000). Human Information Behavior. *Informing Science*, *3*(2), 49–56. Retrieved from http://inform.nu/Articles/Vol3/v3n2p49-56.pdf

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. (1954). *Personality through perception: An experimental and clinical study*. New York: Harper.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the Embedded Figures Test*. Palo Alto, CA: Consulting Psychologists Press.

Yang, Y., & Wu, C. (2009). An attribute-based ant colony system for adaptive learning object recommendation. *Expert Systems with Applications 28*, *36*(2), 3034–3047.

Yuan, F., Liu, B., & Yu, G. (2006). A study on information extraction from PDF files. *Advances in Machine Learning and Cybernetics*, 258–267. Retrieved from http://www.springerlink.com/index/4617511716252121.pdf

Zapalska, A., & Brozik, D. (2006). Learning styles and online education. *Campus-Wide Information System*, *23*(5).

Zipf, G. (1949). *Human Behavior and the Principle of Least-Effort*. Cambridge: Addison-Wesley.

APPENDIX A: QUESTIONNAIRE 1

A SURVEY ON STUDENT'S INFORMATION SEEKING BEHAVIOUR

Thank you for participating in this survey. This survey is being conducted to investigate how postgraduate students seek information for their research. This survey covers variety topics concerning postgraduate student's problem, habit and skills in information seeking.

Information seeking is the process or activity of attempting to obtain information in relation to a goal. It includes recognizing and interpreting the information problem, establishing a plan of search, conducting the search, and evaluating the results.

Notes: Your privacy is considered to be paramount and the information you provide will be treated most confidential and will be used ONLY for the purpose of this research**.** You can withdraw from the experiment at any time. However, your responses to the questions on this short survey, which should take 10-15 minutes to complete, will be taken as an immense contribution towards this research work.

Part 1: Respondent's Background

Instruction: Please tick (/) to answer the following

1. Gender
   - ☐ Male ☐ Female
2. Age
   - ☐ 25 or below ☐ 26 -30 ☐ 31 – 40 ☐ 41 or above
3. Types of Study:
   - ☐ Masters by Coursework
   - ☐ Masters by Research
   - ☐ PhD
4. Research Area:
   - ☐ Artificial Intelligence
   - ☐ Data Communication & Computer Networking
   - ☐ Information Technology
   - ☐ Library and Information Science
   - ☐ Management Information Systems
   - ☐ Multimedia
   - ☐ Software Engineering

5. You are a:
   - ☐ Full-time student ☐ Part-time student

6. Nationality:

Part 2: Investigating student's problem in seeking information

7. Have you received any formal training or orientation on information seeking as part of your academic course?

   ☐ Yes ☐ No

8. Based on the scale provided, mark (/) the one which corresponds with your current behaviour in the course of information seeking.

		Very Often	*Often*	*Sometimes*	*Rarely*	*Almost never*
a.	In resolving information problems I make use of more than one library	☐	☐	☐	☐	☐
b.	The first resource in information retrieval for me is the internet (search engine, subject gateway).	☐	☐	☐	☐	☐
c.	The first resource in information retrieval for me is the library catalog (online, card, on internet).	☐	☐	☐	☐	☐
d.	I am able to think of appropriate keywords for my search.	☐	☐	☐	☐	☐
e.	I am able to connect keyword terms to focus my search.	☐	☐	☐	☐	☐
f.	Information which I have acquired at the beginning of seeking, confirm what I have already known about the problem.	☐	☐	☐	☐	☐
g.	Information seeking has taken me up more time than I presume.	☐	☐	☐	☐	☐

9. Do you face any problem while doing your information seeking process?

   ☐ Yes ☐ No ☐ Not Sure

10. What is the problem you face in information seeking process? Tick all that apply.

- ☐ It is difficult to know where to find relevant information.
- ☐ It is difficult to know how to access the information sources.
- ☐ It is difficult to deal with the large amount of information available.
- ☐ It is difficult to categorize my information needs.
- ☐ It is difficult to understand the information found.
- ☐ It is difficult to ensure that the information is valid.
- ☐ It is difficult to relate the information found to my search subject.
- ☐ Other. (Please specify) ________________________

Part 3: Investigating how students conduct their information seeking

11. Following is a list of information resources, please tick one option, which you use to seek information or you would use to seek information in the future.

		Use it frequently	Use it occasionally	Would use it in the future
a.	Internet search engine (Google, Yahoo)	☐	☐	☐
b.	Library Web Page (OPAC)	☐	☐	☐
c.	Digital library (ACM/IEEE)	☐	☐	☐
d.	Ask Librarian	☐	☐	☐
e.	Experts (a professor/supervisor/lecturer)	☐	☐	☐
f.	Students/ colleagues	☐	☐	☐
g.	Other (Please specify) __________________	☐	☐	☐

12. Please indicate the amount of time you spend per week in information-gathering activities:

- ☐ 0-3 hrs
- ☐ 4-6 hrs
- ☐ 7-9 hrs
- ☐ More than 10 hrs

13. Generally when you try to search for a particular piece of information yourself (i.e. without taking anybody's assistance, guidance, etc.) please state how long it takes to find the needed information?

- ☐ Within a day
- ☐ Within a week
- ☐ Within a month
- ☐ Over a month
- ☐ Difficult to find the needed information without proper assistance/guidance

14. If you cannot find the information at the first attempt, what actions do you take? Tick all that apply.

- ☐ Try another information resources
- ☐ Try another combination keyword
- ☐ Consult expert (i.e. supervisor)
- ☐ Discuss with friends
- ☐ No action
- ☐ Other (s) if any, please specify :____________________

Part 4: Investigating how students evaluate and verify their material

Instruction: Please tick (/) to answer the following

15. Is the information you find normally match your preferences?

☐ Yes ☐ No ☐ Not Sure

16. Do you evaluate the information and sources that you have retrieved?

☐ Yes ☐ No ☐ Do Not Know How

17. How do you identify **the information found is useful to you?** Tick all that apply.

- ☐ It contains a lot of image (graph, diagram) to describe the subject area
- ☐ It has a lot of example to understand the subject area
- ☐ It gives me a lot of definition and explanations
- ☐ It has case study that applies to the subject area
- ☐ Don't know
- ☐ Other (s) if any, please specify :____________________

18. Can you usually complete your research with the information you find?

☐ Yes ☐ No ☐ Not Sure

19. **Please use the space below to add any comments about information seeking:**

Would you be willing to assist the project in the future?

☐ Yes. (please leave your email)

E-mail:

APPENDIX B: QUESTIONNAIRE 2

A Survey on Students' Learning Style

Thank you for participating in this survey. This survey is being conducted to investigate students' awareness of learning style and to identify students' learning style using VARK Questionnaire.

Privacy Statement

Your privacy is considered to be paramount and the information you provide will be treated most confidential and will be used ONLY for the purpose of this research**.** You can withdraw from the experiment at any time. However, your responses to the questions on this short survey, which should take 10-15 minutes to complete, will be taken as an immense contribution towards this research work.

Please select the appropriate response for each question.

Part I: Personal Details

1. What is your gender?

☐ Male

☐ Female

2. Are you an International or a Local student?

☐ International

☐ Local

3. What is the level of your studies?

☐ Masters by Research

☐ PhD

Part II: Learning Style Awareness
Learning Style Definition   For the purposes of this survey, *learning style* is defined as the preference or predisposition of an individual to perceive and process information in a particular way or combination of ways (Zapalska & Brozik 2006).

4. Are you familiar with the term *learning style* before this survey?

☐ Yes

☐ No

If your answer is **<u>Yes</u>** then go to Question 6.

If your answer is **<u>No</u>** then go to Question 8.

5. Are you aware what your learning style is?

☐ Yes

☐ No

6. Do you think knowing your own learning style is important in improving your learning ability?

☐ Yes

☐ No

Learning Style Preference   *Based on VARK Learning Style (Flemings 1992), there are four types of preferences:*
▪ **Visual (V) -** Visual learners prefer the use of diagrams, pictures, videos, slides, graphs and flow charts to represent printed information.   ▪ **Aural / Auditory (A) -** This perceptual mode describes a preference for information that is "heard or spoken."   ▪ **Read/write (R) -** Read-write learners prefer printed words and text as a means of taking in information.   ▪ **Kinaesthetic (K) -** Kinaesthetic preference refers to learning achieved through the use of experience and practice.

7. In your opinion, what is your strongest preference when learning (refer table above)?

☐ Visual

☐ Aural/Auditory

☐ Read/Write

☐ Kinesthetic

Part III: Comments

8. Please use the space below to add any comments about learning style:

Would you be willing to assist the project in the future?

- [ ] Yes. (please leave your email)

E-mail:

Part IV: Learning Style Test
This is VARK Questionnaire (Version 7.1).

Choose the answer which best explains your preference and circle the letter(s) next to it. **Please circle more than one** if a single answer does not match your perception. Leave blank any question that does not apply.

1. You are helping someone who wants to go to your airport, town centre or railway station. You would:
A. go with her.
B. tell her the directions.
C. write down the directions.
D. draw, or give her a map.

2. You are not sure whether a word should be spelled `dependent' or `dependant'. You would:
A. see the words in your mind and choose by the way they look.
B. think about how each word sounds and choose one.
C. find it in a dictionary.
D. write both words on paper and choose one.

3. You are planning a holiday for a group. You want some feedback from them about the plan. You would:
A. describe some of the highlights.
B. use a map or website to show them the places.
C. give them a copy of the printed itinerary.
D. phone, text or email them.

4. You are going to cook something as a special treat for your family. You would:
A. cook something you know without the need for instructions.
B. ask friends for suggestions.
C. look through the cookbook for ideas from the pictures.
D. use a cookbook where you know there is a good recipe.

5. A group of tourists want to learn about the parks or wildlife reserves in your area. You would:
A. talk about, or arrange a talk for them about parks or wildlife reserves.
B. show them internet pictures, photographs or picture books.
C. take them to a park or wildlife reserve and walk with them.
D. give them a book or pamphlets about the parks or wildlife reserves.

6. You are about to purchase a digital camera or mobile phone. Other than price, what would most influence your decision?
A. Trying or testing it.
B. Reading the details about its features.
C. It is a modern design and looks good.
D. The salesperson telling me about its features.

7. Remember a time when you learned how to do something new. Try to avoid choosing a physical skill, eg. riding a bike. You learned best by:
A. watching a demonstration.
B. listening to somebody explaining it and asking questions.
C. diagrams and charts - visual clues.
D. written instructions – e.g. a manual or textbook.

8. You have a problem with your heart. You would prefer that the doctor:
A. gave you a something to read to explain what was wrong.
B. used a plastic model to show what was wrong.
C. described what was wrong.
D. showed you a diagram of what was wrong.

9. You want to learn a new program, skill or game on a computer. You would:
A. read the written instructions that came with the program.
B. talk with people who know about the program.
C. use the controls or keyboard.
D. follow the diagrams in the book that came with it.

10. I like websites that have:
A things I can click on, shift or try.
B interesting design and visual features.
C interesting written descriptions, lists and explanations.
D audio channels where I can hear music, radio programs or interviews.

11. Other than price, what would most influence your decision to buy a new non-fiction book?
A The way it looks is appealing.
B Quickly reading parts of it.
C A friend talks about it and recommends it.
D It has real-life stories, experiences and examples.

12 You are using a book, CD or website to learn how to take photos with your new digital camera. You would like to have:
A a chance to ask questions and talk about the camera and its features.
B . clear written instructions with lists and bullet points about what to do.
C diagrams showing the camera and what each part does.
D many examples of good and poor photos and how to improve them.

13. Do you prefer a teacher or a presenter who uses:
A demonstrations, models or practical sessions.
B question and answer, talk, group discussion, or guest speakers.
C handouts, books, or readings.
D diagrams, charts or graphs.

14. You have finished a competition or test and would like some feedback. You would like to have feedback:
A using examples from what you have done.
B using a written description of your results.
C from somebody who talks it through with you.
D using graphs showing what you had achieved.

15. You are going to choose food at a restaurant or cafe. You would:
A choose something that you have had there before.
B listen to the waiter or ask friends to recommend choices.
C choose from the descriptions in the menu.
D look at what others are eating or look at pictures of each dish.

16. You have to make an important speech at a conference or special occasion. You would:
A make diagrams or get graphs to help explain things.
B write a few key words and practice saying your speech over and over.
C write out your speech and learn from reading it over several times.
D gather many examples and stories to make the talk real and practical.

Thank You for Your Time

APPENDIX C: LIST OF RM

Table 1: List of RM

No	ID	Title	Author	Year	Pages	Topic
	AI 3	Introduction Artificial Intelligence	RC Chakraborty	2010	51	AI
2.	AI 4	Artificial Intelligence: An Overview	Vasant Honavar	2006	14	AI
3.	AI 5	Defining Artificial Intelligence	David B. Fogel	2006	32	AI
4.	AI 6	Artificial Intelligence			62	AI
5.	AI 7	Steps Toward Artificial Intelligence	Marvin Minsky	1960	33	AI
6.	AI 9	Developing an Artificial Intelligence Engine	Michael van Lent & John Laird		11	AI
7.	AI 10	The History of Artificial Intelligence	Chris Smith	2006	27	AI
8.	AI 11	A (Very) Brief History of Artificial Intelligence	Bruce G. Buchanan	2005	8	AI
9.	AI 13	Overview and Tutorial on Artificial Intelligence Systems	Eric Conrad		6	AI
10.	DM1	Introduction to Data Mining	Osmar R. Zaïane	1999	15	DM
11.	DM3	Introduction to Data Mining and Knowledge Discovery		1999	40	DM
12.	DM5	An Introduction to Data Mining	Kurt Thearling		13	DM
13.	DM6	Data Mining: An Overview	Jeffrey W. Seifert	2004	19	DM
14.	DM7	Data Mining	K.U.Leuven		51	DM
15.	DM13	Data-Mining Concepts			18	DM
16.	e-learning 1	e-Learning The Future of Learning	David Wilson	2000	34	e-learning
17.	e-learning 2	Best Practices for Creating E-Learning	Rose Jorgensen	2006	34	e-learning
18.	e-learning 3	Effective Practice with e-Learning	Sarah Knight	2004	60	e-learning
19.	e-learning 4	A Guide to e-learning 1.0	Chris Hall, Nicola van den Berg & Kemi Adamson	2007	8	e-learning
20.	e-learning 5	Learning Objects and E-Learning: an Informing Science Perspective	Eli B. Cohen	2006	12	e-learning
21.	e-learning 6	Practical Applications of Technology for Learning	Brent Schlenker	2008	8	e-learning
22.	e-learning 7	What is E-Learning?	NA	NA	13	e-learning
23.	e-learning 8	e-Learning Series –A Guide for Teachers	Allison Littlejohn & Carol Higgison	2003	35	e-learning
24.	e-learning 9	A Guide to E-Learning: Towards a Knowledge Enterprise	Lesley Mackenzie-Robb	2004	13	e-learning
25.	e-learning 10	E-Learning	Jeremy Francis	2012	6	e-learning
26.	IT 3	Information Technology: Its Impact on Undergraduate Education in Science, Mathematics, Engineering and Technology			55	IT

27.	IT 4	Information Technology in the Classroom		2004	13	IT
28.	IT 6	Introduction: Understanding Information Technology			10	IT
29.	IT 7	Introduction to Information & Communications Technology		2000	27	IT
30.	IT 8	Information and Communication Technology		2007	146	IT
31.	IT 11	How Fluent with Information Technology are our Students?	Sharon Fass McEuen	2001	10	IT
32.	KM1	Doing Knowledge Management	Joseph M. Firestone & Mark W. McElroy	2004	29	KM
33.	KM3	What is Knowledge Management?			16	KM
34.	KM4	Knowledge Management Dealing Intelligently With Knowledge	Drs. Rob van der Spek & Dr. André Spijkervet	2005	60	KM
35.	KM5	Knowledge Management In Education: Defining The Landscape	Lisa A. Petrides & Thad R. Nodine	2003	34	KM
36.	KM6	Introduction to Knowledge Management	Jean Louis Ermine		12	KM
37.	KM7	troduction To Knowledge Management Marifish Era-Net	Dr Alister Scott	2008	15	KM
38.	KM10	Knowledge Communication Problems between Experts and Mangers	Martin J. Eppler	2004	28	KM
39.	Network1	Networking Concepts – Skills for the Electronic Workplace	Stephen Carr	1998	13	Network
40.	Network2	Computer Networking Technologies and Application to IT Enabled Services	Pranab Kumar Chakravarty		11	Network
41.	Network3	Networking Concepts			18	Network
42.	Network4	Networking Basics		2002	10	Network
43.	RM2	Educational research: some basic concepts and terminology	T. Neville Postlethwaite	2005	57	RM
44.	RM5	Research Methodology and Design			44	RM
45.	RM6	Basic Concepts in Research and Data Analysis			20	RM
46.	RM7	Paradigms, Methodology & Methods	Dr. Shelley Kinash		7	RM
47.	RM9	Research and the Research Problem		2011	56	RM
48.	RM12	Research Methodology	Getu Degu & Tegbar Yigzaw	2006	138	RM
49.	RM13	Business Research Methods	Dr. Sue Greener	2008	110	RM
50.	RM16	Research Methodology			41	RM
51.	RM17	Principles and Methods of Development Research	Jan van den Akker		14	RM
52.	System1	A System Development Method	M. A. Jackson	1982	13	System
53.	System2	System Development Life Cycle Guide		2008	50	System

54.	System3	Design of Web-based Information Systems – New Challenges for Systems Development?	Peter H.Carstensen & Lasse Vogelsang	2001	12	System
55.	System5	System Development Methodology	Leslie Williams	2006	35	System
56.	System6	Information Systems Development Techniques and Their Application to do the Hydrologic Database Derivation Application	Paul Davidson		11	System
57.	System7	Information Systems Development Methodologies in the Age of Digital Economy	Dimitrios Papatsoutsos		28	System
58.	System9	An Overview of Systems Design and Development Methodologies with Regard to the Involvement of Users and Other Stakeholders	Shawren Singh & Paula Kotzé	2003	11	System
59.	System10	System Development Methodologies			10	System
60.	Thesis1	MEd Thesis and Project Guidelines		2010	26	Thesis
61.	Thesis2	Manual for Writing a Thesis		2011	15	Thesis
62.	Thesis3	Guide to Theses and Dissertations			39	Thesis
63.	Thesis4	Writing A Thesis Proposal: Independent Learning Resources	Henrike Korner & Helen Drury	2001	47	Thesis
64.	Thesis5	How Do I Successfully Write a Masters Thesis?	Dr. Anna Kreikemeyer & Dr. Patricia Schneider	2007	22	Thesis
65.	Thesis6	How to Write A Strong Thesis Statement - A Writing Centre Handout	Y-Dang Troeung		8	Thesis
66.	Thesis7	How to Write a Thesis: A Working Guide	R Chandrasekhar	2002	33	Thesis
67.	Thesis8	Introductions and Thesis Statements	Williams, Joseph M. & Gregory G. Colomb	2003	4	Thesis
68.	Thesis9	Introduction to Thesis Writing – Structures and Processes	Pam Mort	2002	24	Thesis
69.	Thesis10	Guidelines for Writing a Thesis or Dissertation	Linda Childers & Kurt Kent	2008	8	Thesis
70.	Writing 2	Writing Literature Review		2005	9	Writing
71.	Writing 3	Writing a Literature Review	Allyson Skene		2	Writing
72.	Writing 4	A Guide to Writing the Dissertation Literature Review	Justus J. Randolph	2009	13	Writing
73.	Writing 5	Literature Review Tips	Shannon Mattern		8	Writing
74.	Writing 6	Writing Literature Review			8	Writing
75.	Writing 7	Writing a Literature Review		2006	10	Writing
76.	Writing 9	A Guide to Writing the Dissertation Literature Review	Justus Randolph		17	Writing
77.	Writing 10	Writing a Literature Review			3	Writing

APPENDIX D: DATA FOR FEATURE EXTRACTION TESTING

Table 1: *R* value for each RM

RM	Page	Visual	Read/write	Kinesthetic
A	11	11	11	11
B	40	40	40	40
C	13	13	13	13
D	20	20	20	20
E	13	13	13	13
Total		97	97	97

Table 2: Number of feature vectors expected from RM (E)

RM	Page	Visual	Read/write	Kinesthetic
A	11	6	11	6
B	40	13	40	30
C	13	3	13	4
D	20	4	20	17
E	13	11	13	10
Total		37	97	67

Table 3: Number of feature vectors that are found (F)

RM	Page	Visual	Read/write	Kinesthetic
A	11	6	11	6
B	40	11	40	30
C	13	2	13	4
D	20	2	20	17
E	13	11	13	10
Total		32	97	67

APPENDIX E: RM EXTRACTION RESULTS

Table 1: RM Extraction Results

No.	File Name	Feature Vector Values		
		Visual	Read/Write	Kinesthetic
1.	AI 3	9.68 %	82.26 %	8.06 %
2.	AI 4	10.53 %	73.68 %	15.79 %
3.	AI 5	8.51 %	68.09 %	23.40 %
4.	AI 6	22.12 %	58.65 %	19.23 %
5.	AI 7	30.88 %	48.53 %	20.59 %
6.	AI 9	26.09 %	47.82 %	26.09 %
7.	AI 10	2.22 %	60.00 %	37.78 %
8.	AI 11	27.78 %	44.44 %	27.78 %
9.	AI 13	27.27 %	54.55 %	18.18 %
10.	DM1	38.46 %	38.46 %	23.08 %
11.	DM3	15.66 %	48.20 %	36.14 %
12.	DM5	4.35 %	56.52 %	39.13 %
13.	DM6	0.00 %	70.37 %	29.63 %
14.	DM7	50.00 %	50.00 %	0.00 %
15.	DM13	0.00 %	100.00 %	0.00 %
16.	e-learning 1	15.69 %	66.66 %	17.65 %
17.	e-learning 2	42.50 %	42.50 %	15.00 %
18.	e-learning 3	25.38 %	41.54 %	33.08 %
19.	e-learning 4	7.69 %	61.54 %	30.77 %
20.	e-learning 5	15.00 %	60.00 %	25.00 %
21.	e-learning 6	33.33 %	44.45 %	22.22 %
22.	e-learning 7	10.53 %	68.42 %	21.05 %
23.	e-learning 8	26.03 %	47.94 %	47.94 %
24.	e-learning 9	15.00 %	65.00 %	20.00 %
25.	e-learning 10	40.00 %	40.00 %	20.00 %
26.	IT 3	1.43 %	74.28 %	24.29 %
27.	IT 4	0.00 %	59.09 %	40.91 %
28.	IT 6	10.53 %	52.63 %	36.84 %
29.	IT 7	15.79 %	71.05 %	13.16 %
30.	IT 8	8.33 %	67.13 %	24.54 %
31.	IT 11	31.25 %	62.50 %	6.25 %
32.	KM1	17.78 %	64.44 %	17.78 %
33.	KM3	21.21 %	48.49 %	30.30 %
34.	KM4	27.72 %	58.42 %	13.86 %
35.	KM5	10.20 %	67.35 %	22.45 %
36.	KM6	22.22 %	44.45 %	33.33 %
37.	KM7	0.00 %	55.56 %	44.44 %
38.	KM10	31.48 %	51.85 %	16.67 %
39.	Network1	21.43 %	46.43 %	32.14 %
40.	Network2	21.43 %	46.43 %	32.14 %
41.	Network3	4.35 %	73.91 %	21.74 %
42.	Network4	30.00 %	50.00 %	20.00 %
43.	RM2	0.00 %	62.22 %	37.78 %
44.	RM5	9.09 %	80.00 %	10.91 %
45.	RM6	11.90 %	47.62 %	40.48 %
46.	RM7	8.33 %	58.34 %	33.33 %
47.	RM9	10.64 %	59.57 %	29.79 %
48.	RM12	43.53 %	43.54 %	12.93 %

49.	RM13	12.63 %	57.37 %	30.00 %
50.	RM16	3.92 %	80.39 %	15.69 %
51.	RM17	4.35 %	60.87 %	34.78 %
52.	System1	32.35 %	38.24 %	29.41 %
53.	System2	48.08 %	48.07 %	3.85 %
54.	System3	9.09 %	54.55 %	36.36 %
55.	System5	2.50 %	87.50 %	10.00 %
56.	System6	18.75 %	68.75 %	12.50 %
57.	System7	2.63 %	71.05 %	26.32 %
58.	System9	27.27 %	50.00 %	22.73 %
59.	System10	17.65 %	58.82 %	23.53 %
60.	Thesis1	18.42 %	68.42 %	13.16 %
61.	Thesis2	10.53 %	78.94 %	10.53 %
62.	Thesis3	9.09 %	88.64 %	2.27 %
63.	Thesis4	11.29 %	75.81 %	12.90 %
64.	Thesis5	2.94 %	64.71 %	32.35 %
65.	Thesis6	15.38 %	61.54 %	23.08 %
66.	Thesis7	11.54 %	63.46 %	25.00 %
67.	Thesis8	0.00 %	100.00 %	0.00 %
68.	Thesis9	11.11 %	66.67 %	22.22 %
69.	Thesis10	10.00 %	80.00 %	10.00 %
70.	Writing 2	13.33 %	60.00 %	26.67 %
71.	Writing 3	0.00 %	66.67 %	33.33 %
72.	Writing 4	19.23 %	50.00 %	30.77 %
73.	Writing 5	29.41 %	47.06 %	23.53 %
74.	Writing 6	9.09 %	72.73 %	18.18 %
75.	Writing 7	7.14 %	7.143 %	21.43 %
76.	Writing 9	10.34 %	58.63 %	31.03 %
77.	Writing 10	37.50 %	37.50 %	25.00 %

APPENDIX F: INSTRUCTIONS FOR EVALUATION

Thank you for participating in this system evaluation. This system evaluation is being conducted to investigate how students' learning needs for information can be accommodated based on their preferred learning style. Evaluation should be done on how the information is presented in the document, not the content.

Privacy Statement

Your privacy is considered to be paramount and the information you provide will be treated most confidential and will be used ONLY for the purpose of this research**.** You can withdraw from the experiment at any time. However, your responses to this system evaluation, which should take 15-20 minutes to complete, will be taken as an immense contribution towards this research work.

Evaluation Procedure

There are four steps involved for the evaluation of the tool.

Step 1: Search for reading material using Keyword Based Search

Step 2: Search for reading material using LS Based Search

Step 3: Give feedback

Step 4: Answer survey

INSTRUCTION

This LISTST system has been tested to work on WinXP/Win7 with Internet Explorer 8.

1. Go to this URL: http://27.54.118.190/yana/
2. You will be directed to the following page:

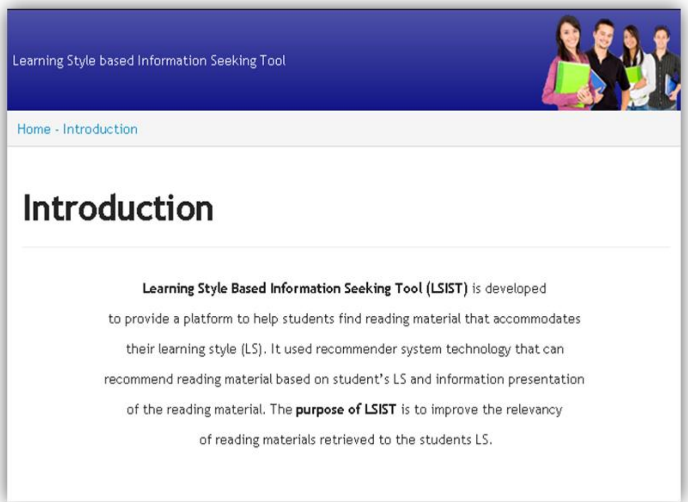


Figure 1: LSIST Homepage

3. Register by clicking register button at the black panel of your top right hand side

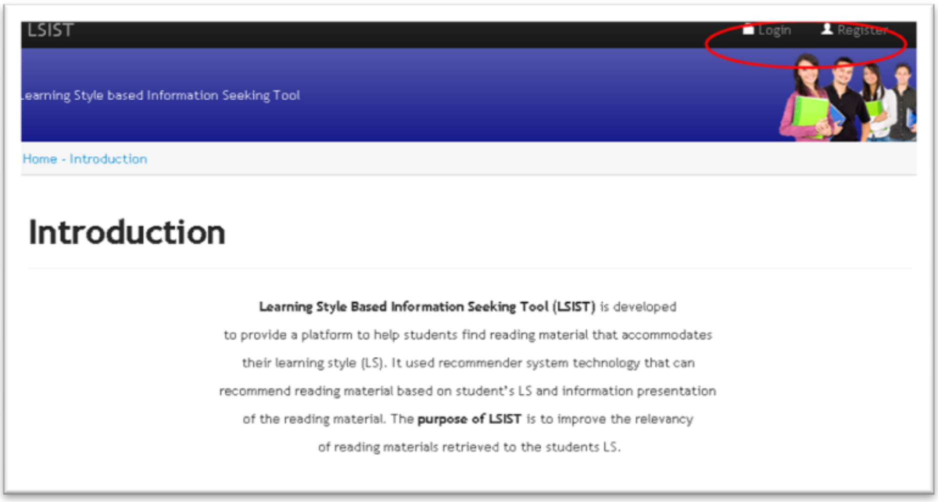


Figure 2: Button register

4. Fill in user detail and click Save

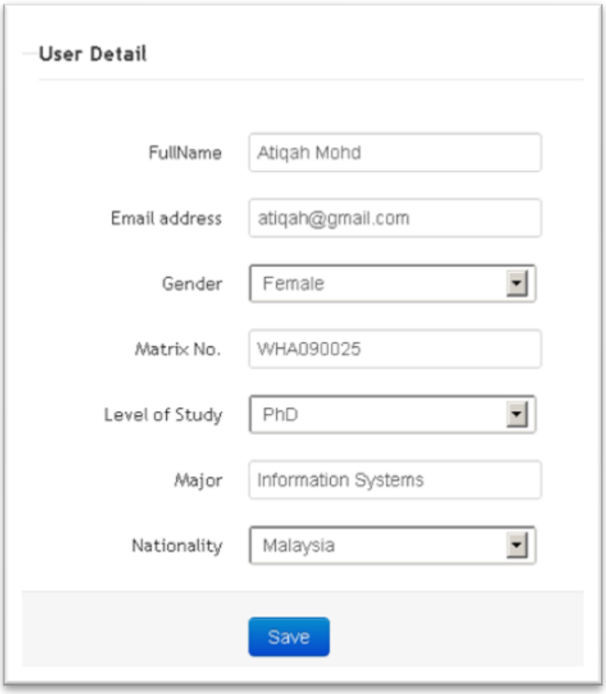


Figure 3: Registration form

5. Then, you will be directed to the Login page as below. Key in your username (email) and password (password).

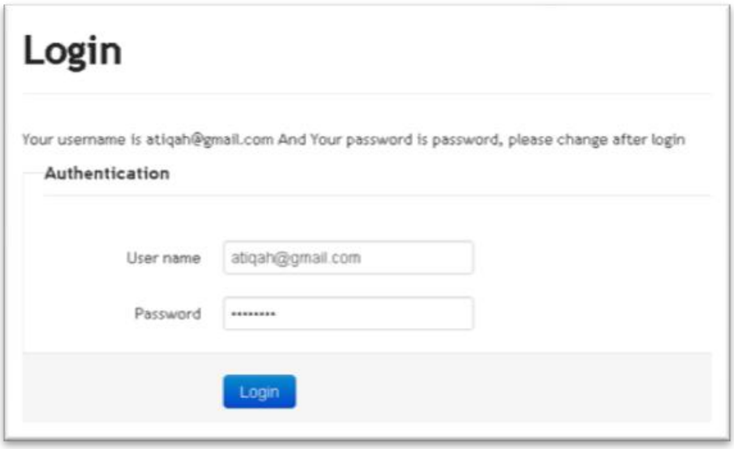


Figure 4: Log in page

6. You will be directed to Student Homepage

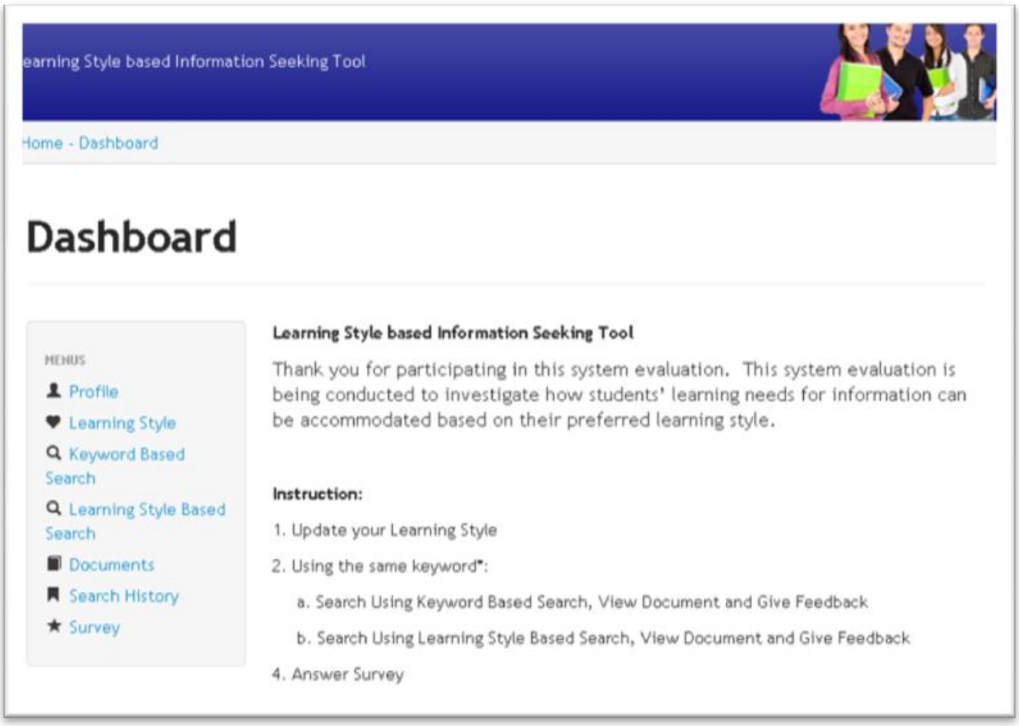


Figure 5: Student Homepage

7. Click tab *Profile* on the dashboard to check and update profile

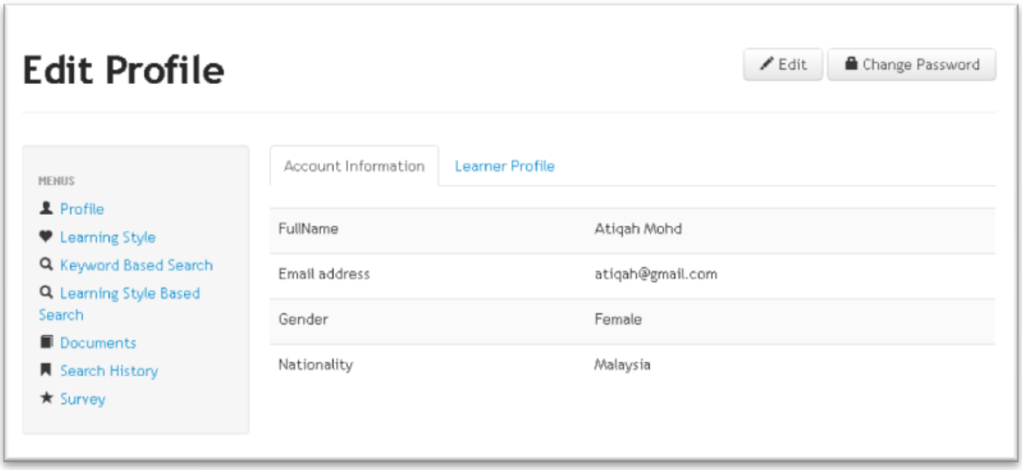


Figure 6: Edit profile page

8. Click tab *Keyword based Search*. Key in keyword and click Search. Suggested keywords: e-learning, data mining, artificial, research methodology, system development, knowledge management

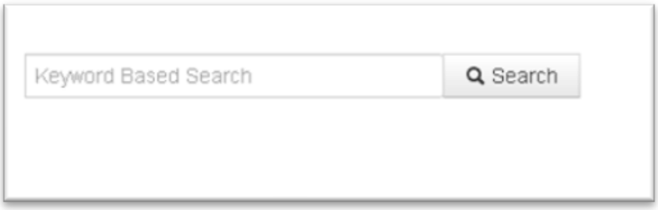


Figure 7: Keyword based Search

9. Results will be displayed as below.

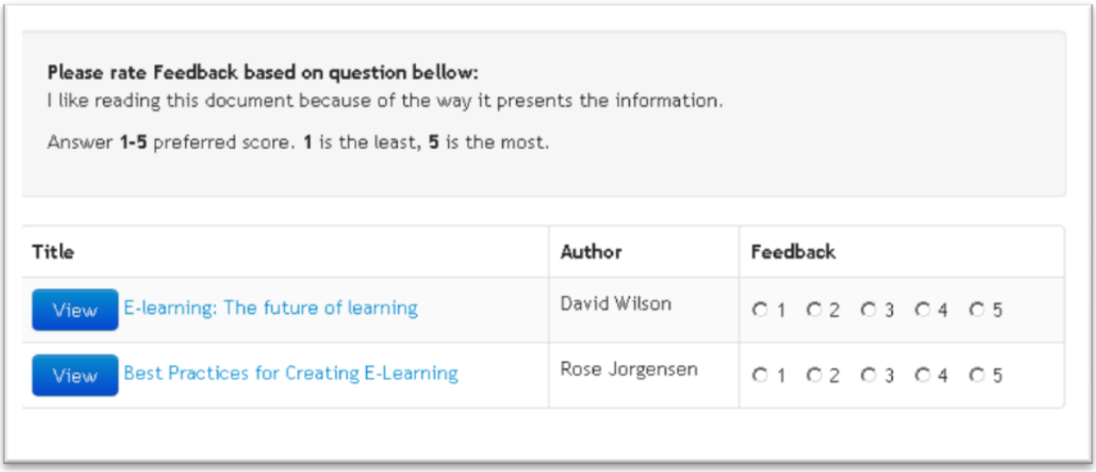


Figure 8: Search results

10. Click button View to view reading material.

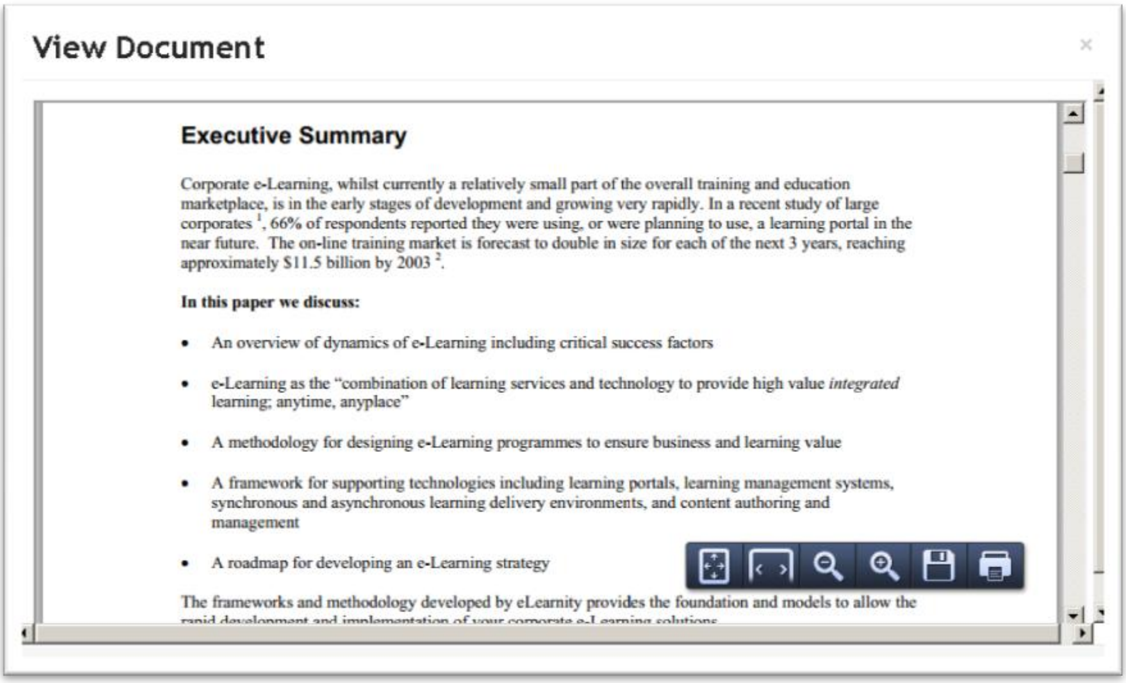


Figure 9: View reading material

11. After reading the material, provide feedback by clicking on the radio button.

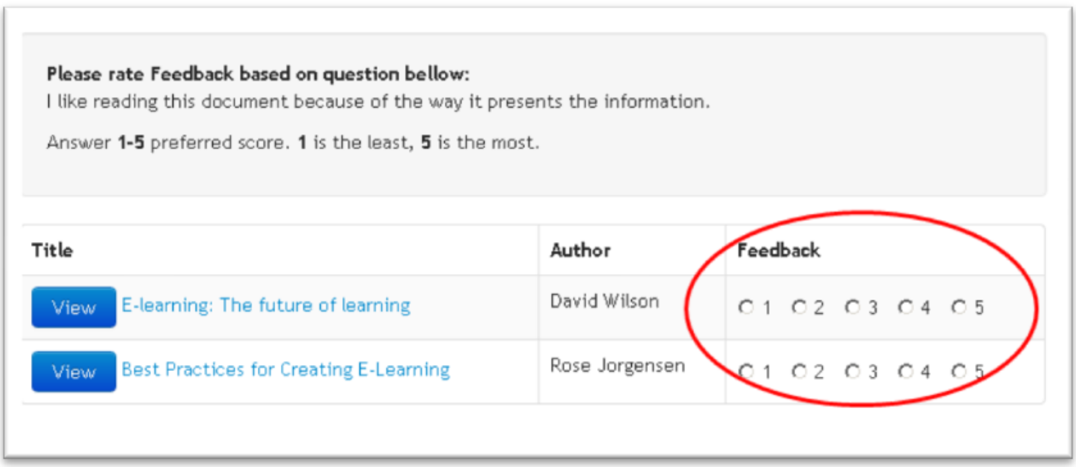


Figure 10: Give feedback

12. Click tab *Learning Style based Search*. Repeat step 8-11.

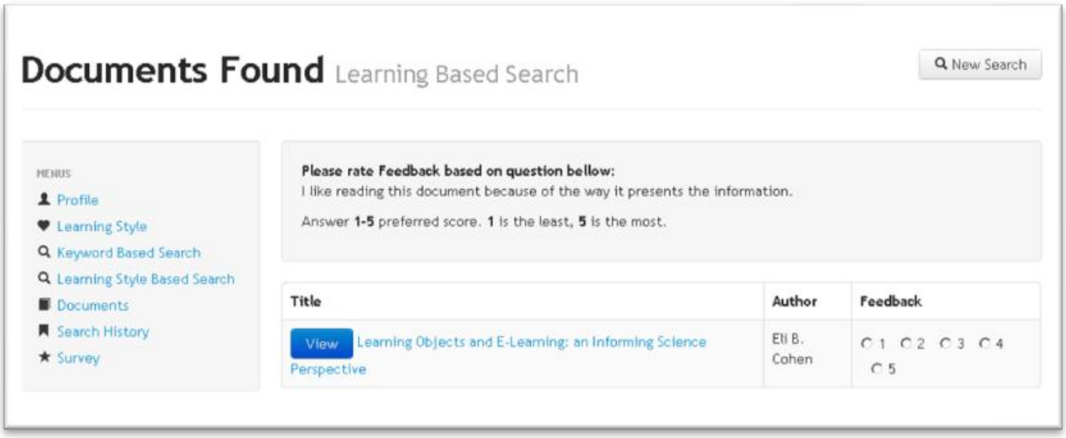


Figure 11: Learning Style based Search

13. Click tab *Survey* and answer Survey

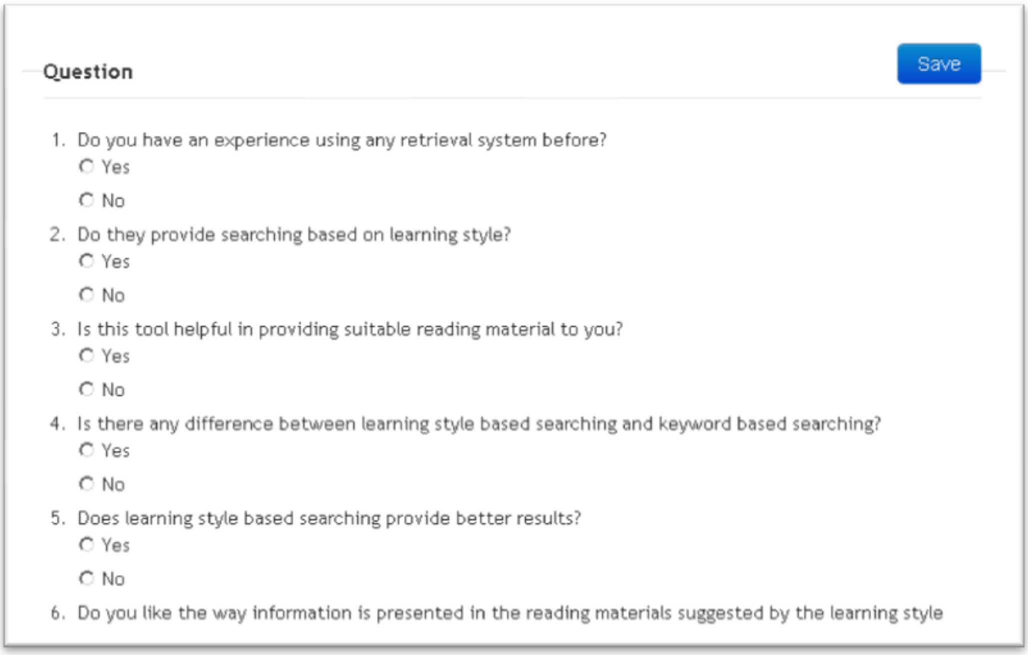


Figure 12: Survey page

14. Click *Save* and you will be directed to Thank You Page.

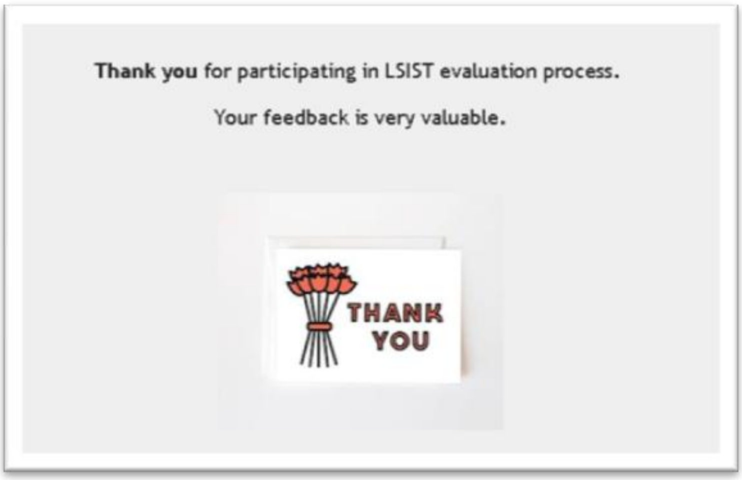


Figure 13: Thank you page

15. Finish.

APPENDIX G: DATA FROM LSIST EVALUATION

Table 1: Evaluation Results

Gender	LevelOf Study	LS	Q1	Q2	Q3	Q4	Q5	PreTest	PostTest
2	1	Aural	1	1	1	1	1	4	5
1	1	Aural	1	1	1	1	2	3	5
2	2	Aural	1	1	1	1	1	2	5
1	2	Aural	1	1	1	1	2	4	3
2	2	Aural	1	1	1	1	1	4	5
2	1	Aural	1	1	1	1	1	3	4
2	2	Aural	1	1	1	1	2	2	4
1	2	Aural	1	1	1	1	1	4	5
2	2	Aural	1	1	1	1	1	3	4
2	2	Aural	1	1	2	1	2	3	3
2	1	Kinesthetic	1	1	1	1	1	4	5
1	2	Kinesthetic	1	1	1	1	1	3	4
2	1	Kinesthetic	1	1	1	1	1	3	4
1	1	Kinesthetic	1	1	1	1	2	4	4
2	1	Kinesthetic	1	1	1	1	1	3	4
1	2	Kinesthetic	1	2	1	1	2	4	4
1	2	Kinesthetic	1	1	1	1	1	3	5
1	2	Kinesthetic	1	1	1	1	2	2	4
2	2	Kinesthetic	1	2	1	1	1	3	4
2	2	Kinesthetic	1	1	1	1	2	3	3
2	2	Multimodal	1	1	2	1	1	2	5
2	1	Multimodal	1	1	1	1	1	4	5
2	2	Multimodal	1	2	1	1	2	4	4
2	1	Multimodal	1	1	1	1	1	3	5
1	2	Multimodal	1	1	1	1	1	3	4
1	2	Multimodal	1	1	1	1	1	3	4
2	2	Multimodal	1	1	1	1	1	3	5
2	1	Multimodal	1	1	1	1	1	2	5
1	2	Multimodal	1	1	1	1	1	3	4
2	2	Multimodal	1	2	1	1	1	4	4
1	1	Read/write	1	1	1	1	1	3	5
2	1	Read/write	1	1	1	1	1	3	4
2	1	Read/write	1	1	1	1	1	3	3
1	1	Read/write	1	1	1	1	1	3	4
2	1	Read/write	1	1	1	1	1	2	3
2	1	Read/write	1	1	1	1	1	3	3
2	1	Read/write	1	1	1	1	1	3	5
1	2	Read/write	1	2	1	1	1	4	4
1	2	Read/write	1	1	2	1	2	3	4

2	2	Read/write	1	1	1	1	1	3	3
2	1	Visual	1	1	1	1	1	3	4
2	1	Visual	2	2	2	1	2	3	4
1	2	Visual	1	2	2	1	2	3	3
2	2	Visual	1	1	1	1	1	2	5
2	2	Visual	1	1	1	1	1	2	4
1	2	Visual	1	1	1	1	2	4	5
2	1	Visual	1	2	2	1	2	4	5
1	2	Visual	1	1	1	1	2	4	5
1	2	Visual	1	1	1	1	1	3	3
1	1	Visual	1	1	1	1	1	4	3

LIST OF PUBLICATIONS

Journals

- Nor Liyana, M. S., Noorhidawati, A. How Graduate Students Seek for Information: Convenience or Guaranteed Result? Malaysian Journal of Library & Information Science. (Accepted) (*ISI Cited Publication*)
- Nor Liyana Mohd Shuib, Rukaini Abdullah. The Architecture of A Learning Style Based Information Seeking Tool, International Journal of Digital Content Technology and its Applications. (Accepted) (*SCOPUS Cited Publication*)
- Nor Liyana Mohd Shuib, Rukaini Abdullah. Integrating Learning Style in Information Seeking Tool. Behaviour & Information Technology. (Submitted) (*ISI Cited Publication)*
- Nor Liyana Mohd Shuib, Rukaini Abdullah. A Learning Style-based Information Seeking Tool Architecture: Design and Evaluation. Computer & Education. (Final draft) (*ISI Cited Publication*)

Conferences

- Nor Liyana Mohd Shuib, Rukaini Abdullah. (2013). LSIST: Learning Style Based Information Seeking Tool. In 8th International Conference on Intelligent Information Processing, 1-3 April 2013, Seoul, Republic of Korea *(Non-ISI/Non-SCOPUS Cited Publication)*
- Nor Liyana Mohd Shuib. Towards adaptive information seeking tool based on learning styles in Post Graduate Research Excellence Symposium (PGReS) 2011, Hotel Blue Wave, Shah Alam, Selangor. *(Non-ISI/Non-SCOPUS Cited Publication)*

- Nor Liyana Mohd Shuib, Noorhidawati Abdullah, Mohammad Hafiz Bin Ismail. *The Use of Information Retrieval Tools: a Study of Computer Science Postgraduate Students*. In International Conference on Science and Social Research (CSSR 2010), 5-7 December, Seri Pacific Hotel, Kuala Lumpur, Malaysia. (*ISI Cited Publication*)
- M. S. Nor Liyana and A. Noorhidawati. Ascertain the Information Seeking Behavior of Computer Science Students. In International Conference on Libraries, Information and Society (ICoLIS 2010), 9-10 November, Petaling Jaya, Malaysia. *(Non-ISI/Non-SCOPUS Cited Publication)*